



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

ÚSTAV AUTOMATIZACE A INFORMATIKY

INSTITUTE OF AUTOMATION AND COMPUTER SCIENCE

ANALÝZA CHEMICKÝCH VAZEB VE SLITINÁCH POMOCÍ DOLOVÁNÍ DAT

Data mining analysis of chemical bonds in alloys

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Vendula Nechutová

VEDOUcí PRÁCE

SUPERVISOR

Ing. Jan Roupec Ph.D.

BRNO 2019

Zadání diplomové práce

Ústav: Ústav automatizace a informatiky
Studentka: **Bc. Vendula Nechutová**
Studijní program: Strojní inženýrství
Studijní obor: Aplikovaná informatika a řízení
Vedoucí práce: **doc. Ing. Jan Roupec, Ph.D.**
Akademický rok: 2018/19

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma diplomové práce:

Analýza chemických vazeb ve slitinách pomocí dolování dat

Stručná charakteristika problematiky úkolu:

Podstata chemických vazeb v pevných látkách, např. slitinách kovů a nekovů, může být teoreticky studována s použitím kvantově–mechanických výpočtů. K posouzení a porovnání interakcí mezi jednotlivými atomy se obzvlášť hodí křivky DOS (density of states) a COHP (crystal orbital Hamilton population) popisující vazebné, nevazebné a protivazebné příspěvky k chemické vazbě. Provést analýzu pro ideální krystal obsahující pouze několik typů interakcí je poměrně jednoduché. V případě planárních defektů, jako jsou například hranice zrn, je situace mnohem složitější kvůli velkému množství odlišných interakcí v okolí defektu v důsledku lokální změny struktury. Takovéto defekty však hrají klíčovou roli pro makroskopické vlastnosti reálných pevných látek, jako jsou mechanické vlastnosti. K pochopení stability hranic zrn je vhodné použít postupy pro dolování dat, kterými je možné odhalit případné posílení či oslabení konkrétních chemických vazeb v jejich okolí.

Cíle diplomové práce:

Cílem práce je analýza dostupného datového souboru DOS a COHP z kvantově–mechanických výpočtů pro různé typy hranic zrn ve slitině Ni₃Si pomocí dolování dat. Výsledkem analýzy bude odhalení takových chemických vazeb v okolí hranic zrn, které způsobují u některých z nich nestabilitu. Samotná analýza bude provedena vlastním programem pro dolování dat, jehož vytvoření je také jedním z cílů práce.

Seznam doporučené literatury:

DRONSKOWSKI, R. Computational Chemistry of Solid State Materials, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005.

HAN, J. Data Mining: Concepts and Techniques, ACADEMIC PRESS, San Diego, 2001.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2018/19

V Brně, dne

L. S.

doc. Ing. Radomil Matoušek, Ph.D.
ředitel ústavu

doc. Ing. Jaroslav Katolický, Ph.D.
děkan fakulty

ABSTRAKT

Práce se zabývá aplikací metod Data miningu pro analýzu dvou mřížek Ni_3Si , jedna se stabilní hranicí zrn a druhá s nestabilní hranicí zrn. Zkoumají se zde křivky DOS a COHP pomocí vybraných metod pro porovnávání křivek. Okolí jednotlivých atomů se vyšetřuje pomocí Voroného diagramu. Tyto informace se pak využijí k odhalení rozdílů ve vazbách mezi stabilní a nestabilní mřížkou.

ABSTRACT

The thesis deals with application of data mining methods for the analysis of two Ni_3Si supercells, one with a stable grain boundary and the second one with unstable grain boundary. DOS and COHP curves are examined using selected curve matching methods. The surroundings of the individual atoms are examined by the Voronoi diagram. This information was used to reveal the differences in binding between stable and unstable supercell.

KLÍČOVÁ SLOVA

DOS, COHP, DTW, Voroného diagramy, porovnání křivek, data mining, odlehlé hodnoty, PCA, Ni_3Si

KEYWORDS

DOS, COHP, DTW, Voronoi diagram, curve comparing, data mining, outlier, PCA, Ni_3Si

BIBLIOGRAFICKÁ CITACE

NECHUTOVÁ, Vendula. Analýza chemických vazeb ve slitinách pomocí dolování dat. Brno, 2019. Dostupné také z: <https://www.vutbr.cz/studenti/zav-prace/detail/117379>. Diplomová práce. Vysoké učení technické v Brně, Fakulta strojního inženýrství, Ústav automatizace a informatiky. Vedoucí práce Jan Roupec.

PODĚKOVÁNÍ

V první řadě bych ráda poděkovala svému vedoucímu práce Ing. Janu Roupcovi Ph.D. za vedení mé práce. Poděkování patří také Ing. Martinu Zelenému Ph.D., který mě k tomuto zajímavému tématu přivedl a konzultoval se mnou v průběhu zpracování problematiku z materiálového hlediska. Nelze opomenout ani cenné rady Ing. Ondřeje Adama a podporu mé rodiny.

ČESTNÉ PROHLÁŠENÍ

Prohlašuji, že tato práce je mým původním dílem, zpracoval jsem ji samostatně pod vedením Ing. Jana Roupce Ph.D. a s použitím literatury uvedené v seznamu literatury.

V Brně dne 1. 3. 2019

.....

Vendula Nechutová

OBSAH

1	ÚVOD.....	15
2	MATERIÁLOVÝ POPIS PROBLEMATIKY	17
2.1	Hranice zrn	17
2.2	Problematika hranic zrn Ni_3Si	18
2.3	Ab initio	18
2.4	DOS	18
2.5	COHP.....	19
2.6	Stabilní a nestabilní mřížka	19
3	POUŽITÉ NÁSTROJE	21
3.1	Programovací jazyk R	21
3.2	RStudio	21
3.3	Nejdůležitější použité knihovny	21
3.3.1	Shiny.....	21
3.3.2	Plotly.....	22
4	POROVNÁNÍ KŘIVEK.....	23
4.1	Dynamic time warping (DTW).....	23
4.2	Detekce odlehlých křivek pomocí horizontálních vrstevnic	26
4.3	Analýza hlavních komponent (PCA).....	27
4.4	Mahalanobisova vzdálenost (Mahalanobis distance)	28
5	VORONÉHO DIAGRAMY	31
5.1	Voroného diagram	31
5.2	Rovina mezi body	32
5.3	Průnik rovin	33
5.4	Transformace souřadnic	34
5.4.1	Bod.....	34
5.4.2	Vektor	34
5.5	Zobrazení vektoru na rovinu.....	35
5.6	Algoritmus	36
6	ANALÝZA.....	41
6.1	Popis dat	41
6.2	Analýza DOS křivek.....	41
6.2.1	Korelace.....	41
6.2.2	Vzdálenost křivek	43
6.2.3	Vzdálenost pomocí DTW	44
6.2.4	Detekce odlehlých křivek pomocí horizontálních vrstevnic	46
6.3	Analýza COHP křivek	47
6.3.1	Korelace.....	47
6.3.2	Vzdálenost křivek	48
6.3.3	Vzdálenost pomocí DTW	49
6.3.4	Detekce odlehlých křivek pomocí horizontálních vrstevnic	50
6.4	Voroného diagramy	51
6.5	Zobrazení jedné mřížky na druhou	52
7	APLIKACE	55
7.1	Křivková analýza DOS	55

7.1.1	Vizuální kontrola.....	55
7.1.2	Korelace, vzdálenost, DTW	55
7.1.3	Detekce odlehlých křivek pomocí horizontálních vrstevnic	56
7.2	Křivková analýza COHP	56
7.3	Zobrazení výsledků analýz.....	57
8	POROVNÁNÍ MŘÍŽEK.....	59
9	ZÁVĚR.....	65
10	SEZNAM POUŽITÉ LITERATURY	67
11	SEZNAM PŘÍLOH	71

1 ÚVOD

V dnešní době vzniká čím dál větší množství dat. Dalo by se říct, že se v datech topíme. Produkuje je každá organizace od bank až po výrobní závody. Jejich množství někdy bývá enormní a kvalitní zpracování bývá obtížné. To může být příčinou toho, že se tyto soubory v organizacích často pouze hromadí a nejsou dál efektivně zpracovávány.

Většina souborů, které často bez povšimnutí leží v elektronických archivech, může vydat mnoho zajímavých informací. Ty lze pak využít k optimalizaci procesů, zpřesnění diagnózy nebo třeba k usnadnění rozhodování.

Jak s těmito daty ovšem nakládat? Najímat více lidí, kteří by se daty „prokousávali“, se s jejich rostoucím množstvím a s čím dál vyšší cenou lidské síly nevyplatí. V takové chvíli nastupuje Data mining. [1]

Pomáhá efektivněji analyzovat velká data, hledat zajímavé vzory a vyhledat neobvyklé části. Využívá například statistiky, strojového učení nebo i umělé inteligence.

Oboru Data mining se věnuji již třetím rokem a dostala jsem se k němu vcelku náhodou. Byla jsem „ulovena“ v rámci náborové akce firmy Bosch, která se rozhodla využít mých znalostí matematiky (v té době jsem byla studentem oboru Matematického inženýrství) v oddělení Industry 4.0, kde mi byl ještě s jedním studentem přiřazen úkol analyzovat data z výroby právě pomocí Data miningu.

Začátky byly krušné a bez řádného vedení jsem postupně získávala nové znalosti o tomto oboru a postupně mě začalo toto téma pohlcovat. Když jsem si tedy vybírala téma diplomové práce, volba pro mě byla vcelku jasná, jediná proměnná v mých plánech bylo, jaká zajímavá data k analýze zvolit.

Tato práce se věnuje pro toto odvětví analýze vcelku neobvyklé. Jedná se o analýzu chemických vazeb ve slitinách. K porovnání interakcí mezi atomy je využito křivek DOS (density of states) a COHP (crystal orbital Hamilton population), které byly získány z teoretického modelování. Křivky DOS náleží k jednotlivým atomům a křivky COHP popisují interakci mezi dvěma atomy.

Využití Data miningu pro analýzu chemických vazeb v mřížkách je dnes celosvětovým trendem. Jeho metody jsou často schopny poskytovat mnohem větší množství informací než běžné experimentální metody.

Pro analýzu máme k dispozici dva soubory, kde jeden popisuje mřížku Ni_3Si se stabilní hranicí zrn, a druhý s nestabilní hranicí zrn. Účelem práce je zjistit pomocí datové analýzy, které vazby způsobují stabilitu nebo nestabilitu. K porovnání dále využijeme data, která máme pro primitivní buňku krystalické mřížky bez hranice zrna.

2 MATERIÁLOVÝ POPIS PROBLEMATIKY

V této kapitole je stručně popsána řešená problematika z pohledu materiálových věd.

2.1 Hranice zrn

Hranice zrn reprezentují jednu z nejdůležitějších tříd rozšířených defektů krystalové mřížky. Jejich vlastnosti jsou zásadní pro mnoho aspektů pevných látek zahrnujících například jejich makroskopickou pevnost.

Hranice zrn se stávají stále důležitějšími z důvodu šíření technologií poskytujících a využívajících ultrajemnozrné a nanozrnité materiály. V těchto materiálech může role vlastností hranic zrna dominovat ve srovnání s oblastí vnitřku zrna. (zrno je tak malé, že povrch zrna tvoří významnou část zrna samotného). To také znamená, že vhodný design hranic zrn může vést k materiálovým systémům s vlastnostmi značně odlišnými od primární mřížky.

Vlastnosti hranic zrn jsou velmi citlivé na změny složení [2] [3] [4], které mohou být například řízeny termodynamicky. Nečistoty, i ve velmi malých koncentracích (jednotky ppm – particles per milion), mohou drasticky změnit charakteristiky hranic zrn. [5]

Jeden z efektů, který je obvykle spojován se segregací nečistot na hranici zrn, je interkrystalické zkrěhnutí. Ten je doprovázen hlubokým snížením tažnosti a pevnosti. Materiály, které výrazně trpí interkrystalickým lomem a nízkou tažností, jsou například niklové Ni_3X intermetalické sloučeniny s krystalovou mřížkou L_{12} . Tyto slitiny [6] [7] však vykazují velký potenciál pro vysokoteplotní aplikace v korozním prostředí. [8] Jejich kohezní pevnost na hranici zrn klesá s rostoucím rozdílem ve valenci mezi Ni a X atomem a s rostoucí velikostí X atomu v pořadí $\text{Ni}_3\text{Al} > \text{Ni}_3\text{Ga} > \text{Ni}_3\text{Si} > \text{Ni}_3\text{Ge}$. [9]

Jiné vysvětlení, odvozené od chování nečistot v čistých kovech, může být založeno na elektronegativitě atomu X. Čím více je elektronegativní atom X na hranici zrn, tím větší tendence k vytažení náboje z Ni-Ni vazeb na hranici, čímž se sníží kohezní pevnost a podpoří interkrystalický lom. [10]

V případě Ni_3Al může být tento problém vyřešen například přidáním malého množství bóru, který zlepšuje kohezi hranic zrn [11] a mění interkrystalický charakter lomu na transkrystalický. Stejně zlepšení vlastností nicméně nebylo pozorováno po přidání bóru do intermetalické sloučeniny Ni_3Si . Bór mírně zvýšil tažnost Ni_3Si , zatímco mód lomu neovlivnil. [6] [12] Na druhou stranu, vlastnosti hranic zrn mohou být silně vylepšeny přidáním velkého množství titanu, což vede k vysoce uspořádaným $\text{Ni}_3(\text{Si}, \text{Ti})$ slitinám. [13] [14]

2.2 Problematika hranic zrn Ni_3Si

Tato práce se věnuje hranici typu $\Sigma 5(210)$. V Ni_3Si je možné pozorovat několik typů hranic zrn, tento typ byl vybrán pro svoji jednoduchost. Analyzují se zde dvě možné varianty rozhraní, které se liší lokálním chemickým složením.

U obou bylo zjištěno, že ve srovnání se základní mřížkou mají výrazně sníženou smykovou elastickou konstantu (C_{55}). [15] [16] Hranice se stává nestabilní součástí mřížky. Jedna varianta rozhraní má pouze zhoršené elastické vlastnosti, zatímco druhá se ukázala být zcela nestabilní. [17]

2.3 Ab initio

Jedná se o výpočty elektronové struktury z prvních principů. To znamená, že pro výpočty nejsou používány žádné empirické modely, ani se nevyužívá experimentálně naměřených hodnot jako vstupních parametrů. Veškeré výsledky jsou založeny na řešení Schrödingerovy rovnice, díky čemuž je možné získat různé vlastnosti látek i ve stavech, které jsou experimentálně obtížně dosažitelné nebo nedosažitelné.

Analytické řešení Schrödingerovy rovnice pro komplexní mnohočásticové systémy tuhých látek není možné, řešení lze ale dosáhnout pouze numericky. Výrazné zjednodušení pak přináší teorie funkcionálu hustoty (DFT – Density Functional Theory). Tato teorie redukuje komplexnost elektronových interakcí v mnohočásticovém systému na efektivnější jedoelektronové rovnice (Kohn-Shamovy rovnice) definované výměnným a korelačním funkcionálem, který závisí pouze na elektronové hustotě daného systému. Ačkoliv přesná forma tohoto funkcionálu není známá, jeho aproximace nám umožňují velmi přesně popsat mnoho materiálových vlastností.

Volba vhodné aproximace funkcionálu patří k nejdůležitějším parametrům, které ovlivňují ab initio výpočty založené na DFT. Mezi další parametry patří například volba báze jedoelektronových vlnových funkcí, které matematicky popisují chování elektronů a rozložení elektronové hustoty, a volba algoritmu pro řešení Kohn-Shamových rovnic a pro výpočet energií, sil a napětí. Zvolená aproximace a vlnové funkce určují přesnost výpočtu, použitý algoritmus je rozhodující pro efektivitu výpočtu. [18]

Data analyzovaná v této práci (křivky DOS a COHP, viz kapitoly 2.4, 2.5) byla získána právě touto metodou, rovněž tak s jejich pomocí byla zjištěna nestabilita hranice popsaná výše.

2.4 DOS

Hustota stavů (Density of states, DOS) je v podstatě počet různých stavů na určité energetické úrovni, které mohou elektrony obsazovat, tj. počet elektronových stavů na jednotku objemu na jednotku energie.

Hustota stavů má výrazný vliv na výsledné vlastnosti daného systému. Především rozhoduje o jeho celkové stabilitě. Na DOS dále závisí například měrná tepelná kapacita,

magnetická susceptibilita, mechanické vlastnosti nebo transportní jevy vodivých pevných látek. Z výpočtů je možné taktéž určit obecné rozdělení stavů jako funkci energie nebo vzdálenost mezi energetickými pásy v polovodičích. [19]

2.5 COHP

COHP (Crystal Orbital Hamilton Populations) je nástroj pro detekci vzájemných interakcí mezi atomy v pevných látkách (krystalických i amorfních). COHP rozděluje energii pásové struktury na orbitální párové interakce.

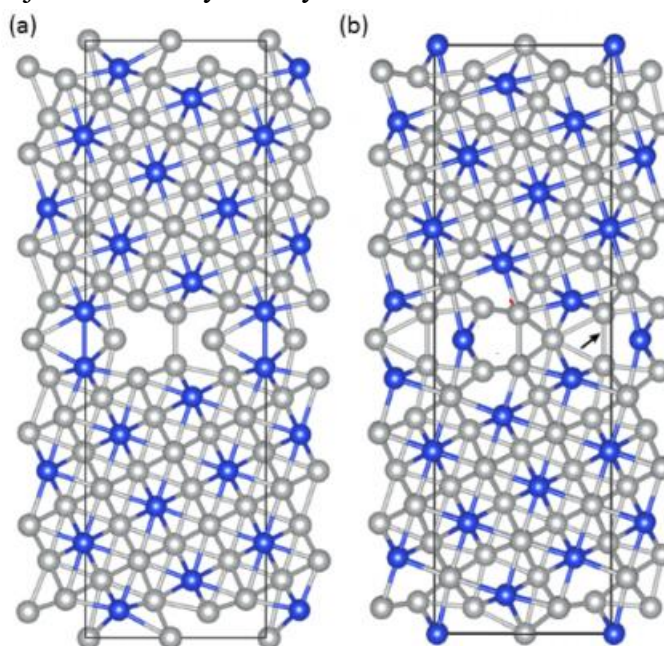
Schéma COHP ukazuje vazební, nevazební a protivazební příspěvky k energii pásové struktury a obvykle se vytváří společně s DOS (které pouze informuje na jaké hladině se elektrony mohou vyskytovat, ale nenese informaci o charakteru jejich vazby).

Stejně jako integrace DOSu udává počet elektronů v systému, energetický integrál COHP ukazuje charakter a sílu chemické vazby. [19]

Z DOS a COHP je možné získat velké množství informací o daném systému. Tato data jsou nicméně natolik obsáhlá, že je běžný způsob jejich zpracování je velmi náročný. Pro nalezení zajímavých závislostí je tedy vhodné využít Data miningu.

2.6 Stablní a nestablní mřížka

Na obr. 2-1 jsou zobrazeny mřížky Ni_3Si se stabilní a nestablní hranicí zrn.



Obrázek 2-1 Ni_3Si mřížka s hranicí zrn (a) stabilní (b) nestablní [17]

3 POUŽITÉ NÁSTROJE

V této kapitole jsou popsány nástroje použité k analýze.

3.1 Programovací jazyk R

R (logo na obr. 3-1) je programovací jazyk primárně určený pro statistické výpočty a jejich vizualizaci. Jedná se o GNU projekt.

R je open source software, což značně napomáhá k jeho rozvoji. K dispozici jsou nyní již tisíce knihoven (packages), které rozšiřují funkcionality jak vizualizační, tak metodické. Mnoho z nich se týká právě Data miningu (obsahují metody Decision tree, Random forest, ...), proto patří společně s Pythonem ke dvěma nejpoužívanějším nástrojům v tomto oboru. [20]



Obrázek 3-1 Logo R
[20]

3.2 RStudio

Pro práci s jazykem R bývá nejčastěji využíváno prostředí RStudia (logo obr. 3-2). Zahrnuje konzoli, editor, ze kterého můžeme přímo provádět kód, nástroje pro vizualizaci a historii.

Jedná se také o open source software, který lze spustit na Windows, Mac a Linux. Existuje i serverová verze spustitelná přes webový prohlížeč. [21]



Obrázek 3-2 Logo RStudio
[21]

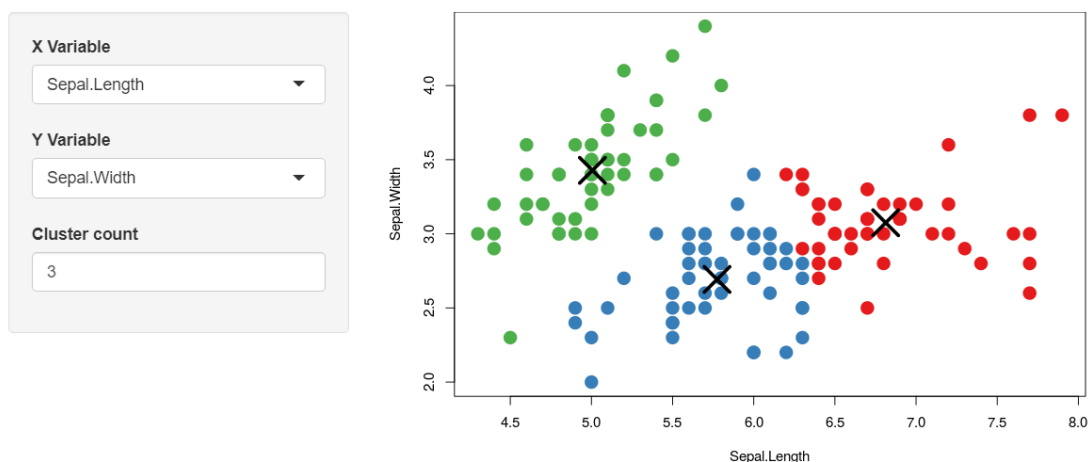
3.3 Nejdůležitější použité knihovny

3.3.1 Shiny

Shiny je knihovnou jazyka R, která pomáhá vytvářet webové aplikace bez nutnosti znalostí HTML, CSS nebo JavaScript. Jejich znalost lze však využít pro vytvoření vlastního designu nebo pro vytvoření pokročilejších funkcionalit.

Vzhledem k zaměření jazyka R se tyto aplikace využívají většinou jako interaktivní dashboardy s možností volby parametrů pomocí různých „widgetů“. Jak taková aplikace může vypadat, je zobrazeno na obr. 3-3. [21]

Iris k-means clustering

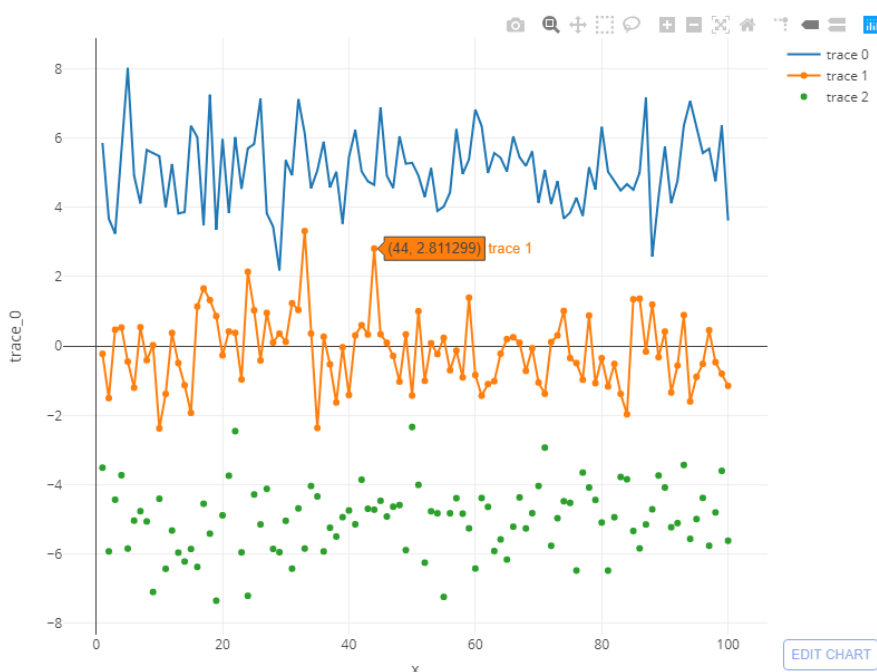


Obrázek 3-3 Ukázka shiny aplikace [21]

3.3.2 Plotly

Tato knihovna vytváří interaktivní a velmi kvalitní grafy. Není dostupná pouze pro R, ale také pro Python, Matlab a další. Umožňuje širokou škálu využití – bodové, sloupcové, liniové, krabicové grafy, heatmaps a další (obr. 3-4). [22]

Přímo v grafu pak máme možnosti přibližování, posouvání, filtrování, zobrazení informací po najetí na určitý bod a stažení grafu ve formátu png. Díky těmto vlastnostem jsou tyto grafy často využívány pro prezentaci. Jejich velkou nevýhodou je však pomalé vykreslování objemnějších dat.



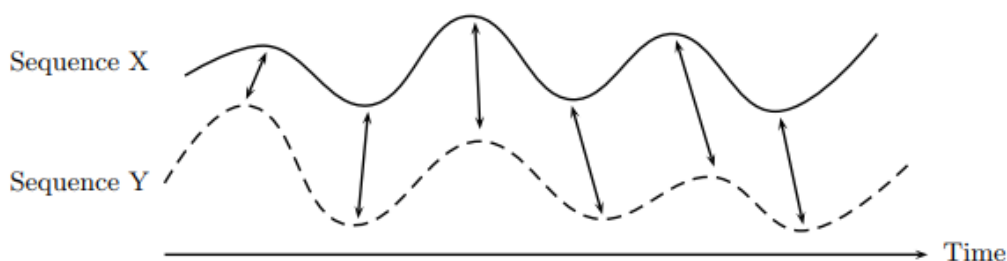
Obrázek 3-4 Ukázka grafu z knihovny plotly [22]

4 POROVNÁNÍ KŘIVEK

Analýza křivek je zatím málo probádaná oblast Data miningu, která stále čeká na nové výzkumy. Algoritmů na analýzu křivek je tedy stále poskrovnu, zde je popsáno několik, které jsou v textu dále využity.

4.1 Dynamic time warping (DTW) [23]

DTW je technika vhodná k hledání optimálního zarovnání dvou časově závislých křivek (obr.4-1). Původně byla vyvinuta za účelem rozpoznávání vzorů v mluveném slově.



Obrázek 4-1 Ilustrace principu DTW [23]

Uvažují se dvě časově závislé sekvence $X = (x_1, x_2, \dots, x_N)$ a $Y = (y_1, y_2, \dots, y_M)$, kde $M, N \in \mathbb{N}$. Sekvence reprezentují diskrétní signály závislé na čase. Tyto sekvence obsahují body z prostoru F

$$x_n y_m \in F \quad n \in [1: N], m \in [1: M]. \quad (1)$$

Pro porovnání dvou bodů z této sekvence se používá míra lokální vzdálenosti, která je definovaná jako funkce

$$c: F \times F \rightarrow \mathbb{R}_{\geq 0}. \quad (2)$$

Pokud si jsou x a y podobné, je $c(x, y)$, velmi malé. V opačném případě je pak $c(x, y)$ vysoké.

Z vypočtených hodnot se vytvoří matice nákladů $C \in \mathbb{R}^{N \times M}$ (cost matrix) tak, že

$$C(n, m) = c(x_m, y_n). \quad (3)$$

Cílem je najít zarovnání, které bude mít co nejnižší náklady. Cesta p je definovaná následovně.

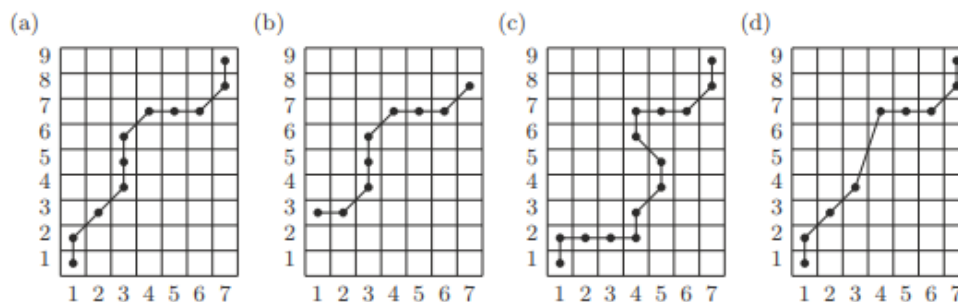
$$p_1 = (n_1, m_1) \in [1: N] \times [1: M] \quad (4)$$

$$p_1 = (1, 1) \quad p_L = (N, M) \quad (i)$$

$$n_1 \leq n_2 \leq \dots \leq n_L \quad m_1 \leq m_2 \leq \dots \leq m_L \quad (ii)$$

$$p_{i+1} - p_i \in \{(1, 0), (0, 1), (1, 1)\} \quad (iii)$$

Podmínka (i) nám určuje začátek a konec cesty p . Monotónnost je zaručena podmínkou (ii). V podmínce (iii) je omezena velikost kroku. Příklad toho, jak může taková cesta vypadat, je na obr. 4-2.



Obrázek 4-2 Ilustrace cest pomocí dvojic indexů dle (2) [23]

- (a) cesta splňující všechny podmínky (i), (ii), (iii)
- (b) cesta porušující podmínku začátku a konce (i)
- (c) cesta porušující monotónnost (ii)
- (d) cesta porušující podmínku velikosti kroku (iii)

Situaci si lze představit jako snahu dostat se přes krajinu, kde hodnoty v matici C představují výšku nad mořem, s co nejmenší námahou. Intuitivně se dá vytušit, že nejsnazší bude jít údolím.

Cesta p definuje zarovnání mezi sekvencemi X a Y přiřazením bodu $x_{n_l} \in X$ k prvku $y_{m_l} \in Y$.

Celkové náklady $c_p(X, Y)$ cesty p mezi X a Y jsou definovány pomocí matice nákladů

$$c_p(X, Y) = \sum_{l=1}^L c(x_{n_l}, y_{m_l}). \quad (5)$$

Optimální cesta p^* mezi X a Y je takové p , pro které jsou celkové náklady (5) minimální. Vzdálenost $DTW(X, Y)$ je definována jako celkové náklady optimální cesty mezi X a Y . (6)

$$DTW(X, Y) = c_{p^*}(X, Y) = \min_p \{c_p(X, Y)\} \quad (6)$$

Pro nalezení optimální cesty by lze projít každou možnou cestu a pak vybrat tu s minimálními celkovými náklady. Tento způsob by byl však výpočetně velmi náročný. K těmto účelům se tedy využívá metod dynamického programování.

Dále bude popsán $O(NM)$ algoritmus. Pro začátek se vytvoří matice D , pro jejíž členy bude platit

$$D(n, m) = DTW(X(1:n), Y(1:m)), \quad (7)$$

kde

$$\begin{aligned} X(1:n) &= (x_1, x_2, \dots, x_n), \\ Y(1:m) &= (y_1, y_2, \dots, y_m). \end{aligned} \quad (8)$$

Matice D se nazývá kumulovanou maticí nákladů. Dále je popsán algoritmus, jak efektivně tuto matici sestojit.

Pro akumulovanou matici nákladů D platí

$$D(n, 1) = \sum_{k=1}^n c(x_k, y_1) \quad \text{pro } n = 1, \dots, N \quad (9)$$

$$D(1, m) = \sum_{k=1}^m c(x_1, y_k) \quad \text{pro } m = 1, \dots, M \quad (10)$$

$$D(n, m) = \min\{D(n-1, m-1), D(n, m-1), D(n-1, m)\} + c(x_n, y_m) \\ \text{pro } 1 < n \leq N, 1 < m \leq M. \quad (11)$$

Výpočet této matice je jednodušší, pokud se přidá nultý řádek a sloupec, jehož hodnoty budou rovny nekonečnu a $D(0,0) = 0$. Potom lze vynechat speciální vzorce pro první sloupec a řádek a využívat pouze poslední (11) se změněnou podmínkou

$$n = 1, \dots, N; \quad m = 1, \dots, M. \quad (12)$$

Vyhledání optimální cesty p^* se začíná od konce

$$p_L = (N, M), \quad (13)$$

Přidávají se postupně další body

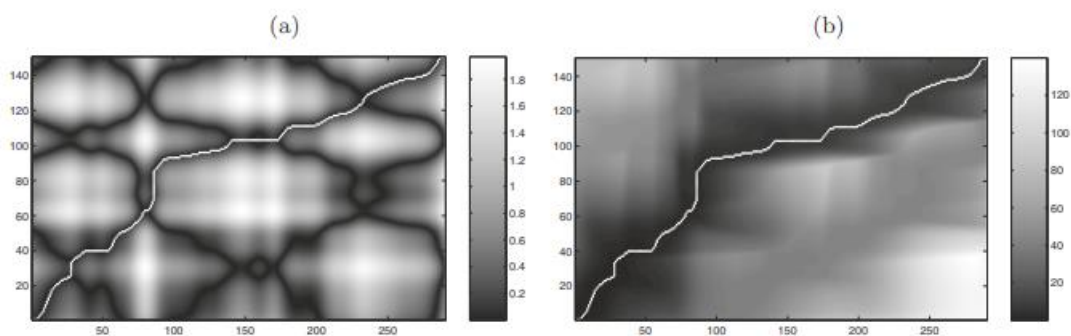
$$p_l = (n, m), \quad (14)$$

dokud nedostaneme $(n, m) = (1, 1)$, v tu chvíli jsme dostali počáteční hodnotu, pro kterou $l = 1$.

Body přidáváme podle následujícího vzorce.

$$p_{l-1} = \begin{cases} (1, m-1) & \text{pro } n = 1 \\ (n-1, 1) & \text{pro } m = 1 \\ \operatorname{argmin}\{D(n-1, m-1), D(m-1, 1), D(1, n-1)\} & \text{jinak} \end{cases} \quad (15)$$

Příklad optimální cesty můžeme vidět graficky znázorněný na obr. 4-3.



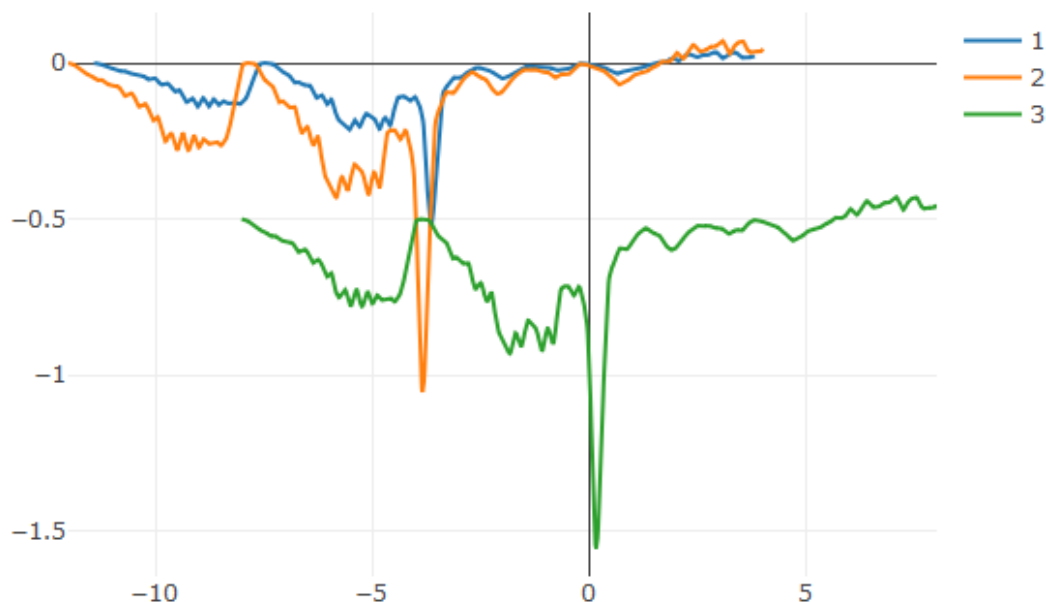
Obrázek 4-3 Optimální cesta p^* přes [23]

(a) matici nákladů C

(b) kumulovaná matice nákladů D

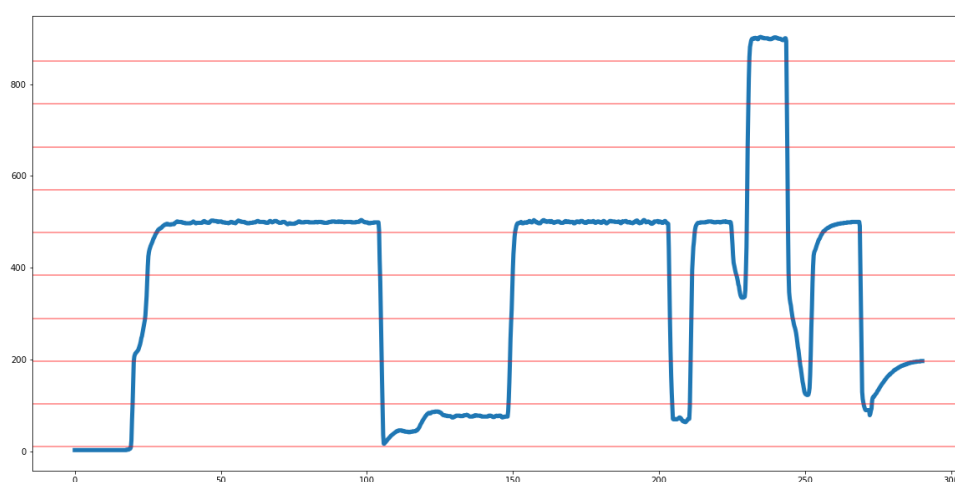
4.2 Detekce odlehlých křivek pomocí horizontálních vrstevnic¹

Tato metoda je vhodná především k hodnocení tvaru křivky bez závislosti na x-ové, případně y-ové ose. To znamená, že není závislá na posunutí křivky, či vynásobení určitou hodnotou. Křivky z grafu z obr. 4-4 by tedy byly vyhodnoceny jako stejné.



Obrázek 4-4 Křivky posunuté a násobené konstantou

Tento algoritmus lze jednoduše popsat následovně. Graf se rozdělí na adaptabilními horizontálními vrstevnicemi (na obr. 4-5 červeně). Tyto vrstevnice rozdělují oblast grafu na stejné části mezi minimem a maximem konkrétní křivky.



Obrázek 4-5 Horizontální vrstevnice

¹ Tento algoritmus byl vytvořen v rámci projektu firmy Bosch, hlavní autor Florian Flatow

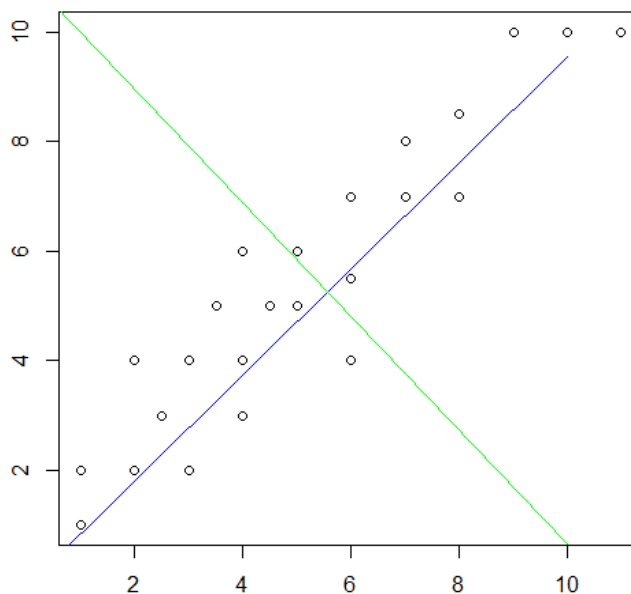
U každé z těchto vrstevnic se pak určí, kolikrát „přetne“ daný graf. Takto lze získat ke každé křivce vektor délky n , který se pak zredukuje pomocí PCA (Principal Component Analysis – Analýza hlavních komponent). Například pro graf na obrázku 4-5 by to byl vektor (2, 2, 2, 2, 10, 10, 8, 9, 7, 1).

Na vektory je možné aplikovat běžnější metody pro vyhledání odlehlých prvků. V tomto algoritmu byla zvolena poměrně jednoduchá varianta, kde se měří vzdálenost od středu shluku, který se získá pomocí jednoduchého průměru všech souřadnic, k jednotlivým bodům pomocí Mahalanobisovy vzdálenosti. Tato hodnota pak určuje „excentricitu“ křivky od ostatních.

4.3 Analýza hlavních komponent (PCA) [24]

Analýza hlavních komponent (Principal component analysis, dále PCA) je metoda vhodná na zdůraznění hlavních rozdílných (případně podobných) rysů u dat s velkým počtem dimenzí. Zároveň může pomoci ke zmenšení objemu dat redukcí dimenzí bez velké ztráty informace.

Pro lepší představu, jak tato metoda pracuje, je zde uveden následující zjednodušený příklad. Máme dvourozměrný dataset bodů vykreslených na obr. 4-6, ale z nějakých důvodů by bylo potřeba používat data pouze o jedné dimenzi.



Obrázek 4-6 Nové osy dle PCA

Prostý výběr buď x -ových či y -ových hodnot by však byl ochuzen o významné informace, proto je nejdříve vhodně transformujeme. Je patrné, že data mají poměrně vysokou korelaci, takže pokud se nová osa bude vést, jak je znázorněno na obr. 4-6 modrou čarou, podchytí se maximální variabilita shluku, zatímco hodnoty, které se objeví na zelené ose, obsahují poměrně málo informací, je tedy možné ji zanedbat.

Jak se ale tyto nové souřadnice dostanou? Ze všeho nejdříve se určí korelační matice zkoumaných dat. K této matici se poté vypočtou vlastní čísla a vlastní vektory.

Vlastní vektory pak mají význam směru nových souřadnic a vlastní čísla určují významnost odpovídající dimenze.

Z vlastních vektorů se vytvoří matice F tak, že do jejích řádků se dosadí vlastní vektory seřazené od nejvíce k nejméně významnému. Nyní je možné udělat závěrečnou transformaci souřadnic

$$X_{\text{nové}}^T = F \times X_{\text{staré}}^T. \quad (16)$$

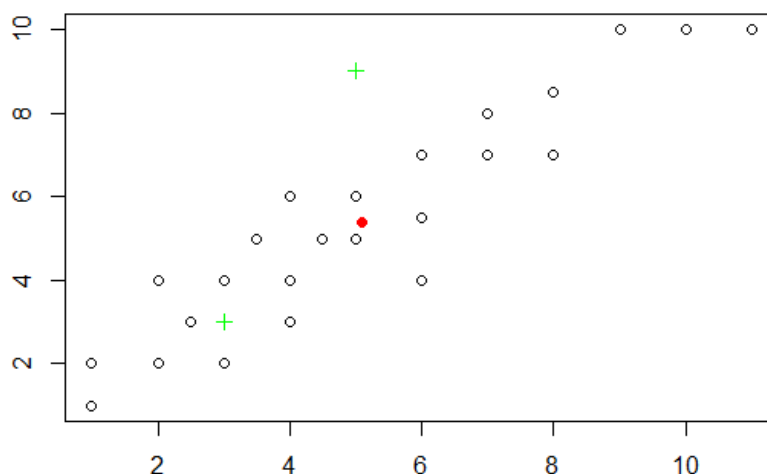
Poté lze rozhodnout dle významnosti dimenze (z vlastních čísel) a některé sloupce vynechat.

4.4 Mahalanobisova vzdálenost (Mahalanobis distance) [25]

Mahalanobisova vzdálenost se zpravidla používá pro hledání odlehlých hodnot. Její výhoda oproti Euklidově je, že dokáže lépe pracovat s korelovanými daty.

Například při pohledu na graf na obr. 4-7, kde jsou černé body vysoce korelované (v tomto případě je korelace 0,93), lze už tušit, že pomocí klasické Euklidovy vzdálenosti se budou odlehlé hodnoty hledat těžko.

Jako příklad lze konkrétně porovnat body znázorněné zelenými křížky, jejichž euklidovská vzdálenost od středu shluku (červený bod) je stejná. Přesto je očividné, že bod (5, 9) méně zapadá mezi ostatní než bod (3, 3).



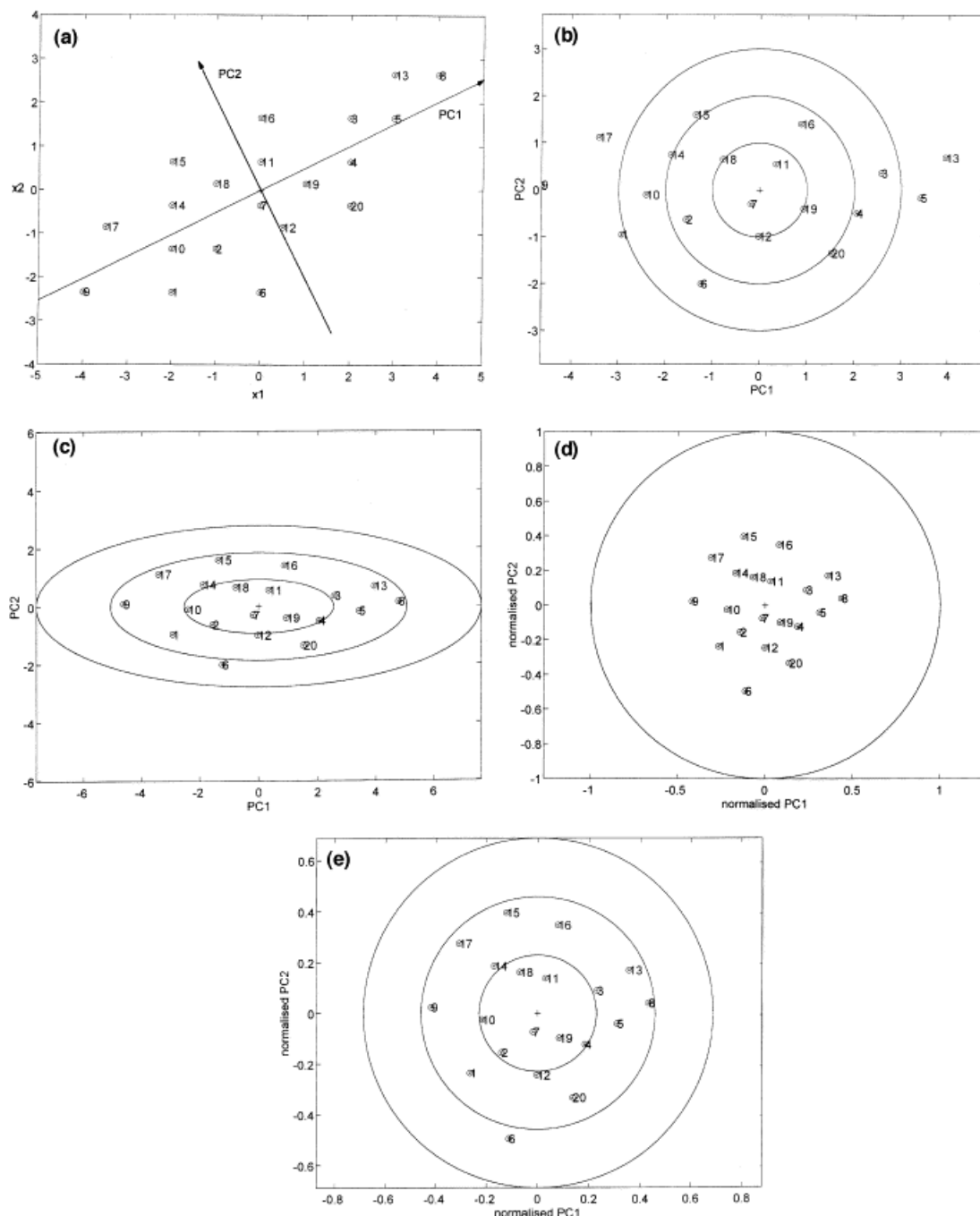
Obrázek 4-7 Korelovaná data

V takových případech je ideální přejít ke vzdálenosti Mahalanobisově. Výpočet si lze představit, jako použití klasické Euklidovy vzdálenosti po transformaci souřadnic, kdy se zvolí, podobně jako v PCA, nové osy. První se zvolí tak, že prezentuje maximální variabilitu bodů, každá další je pak ortogonální na tu první a snaží se zahrnout další maximální variabilitu.

Aby bylo možné použít Euklidovu vzdálenost, musí se osy „zkrátit“ či „prodloužit“ oproti sobě, čímž vznikne shluk bodů s korelací blížíící se nule. Touto

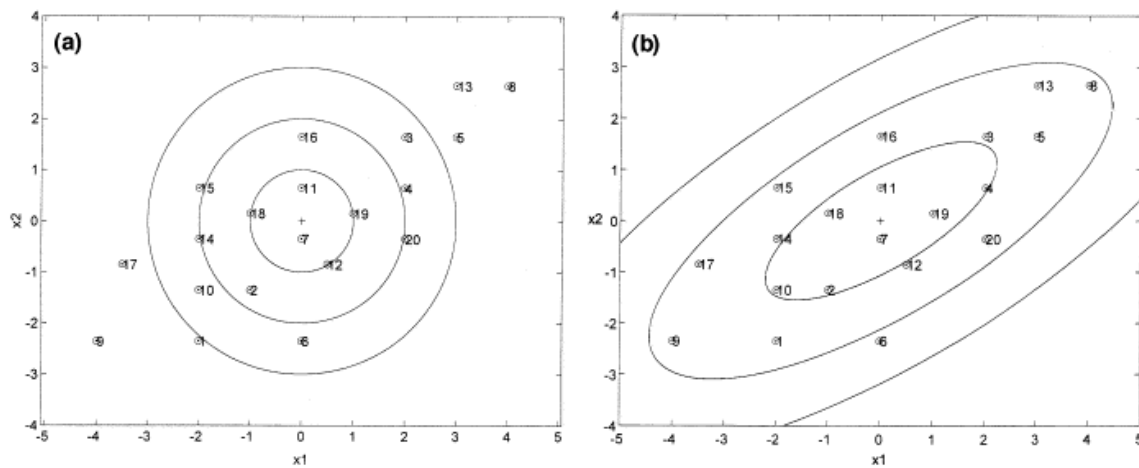
transformací se odstraní problém, který byl nastíněn výše, a může se nyní počítat s klasickou Euklidovskou vzdáleností, jak jsme zvyklí.

Na obr. 4-8 je uveden příklad ve 2D, postup lze však jednoduše aplikovat i pro vyšší dimenze.



Obrázek 4-8 Mahalanobisova vzdálenost – ilustrace pomocí transformace [25]

Jiná přestava pak může být taková, že pokud se znázorní graficky euklidova vzdálenost od středu shluku, pak vrstevnice tohoto grafu by byly soustředné kružnice. U Mahalanobisovi vzdálenosti by to pak byly soustředné elipsy, které mají hlavní osu ve směru největší variability. (obr. 4-9)



Obrázek 4-9 Mahalanobisova vzdálenost – ilustrace pomocí vrstevnic [25]

Matematicky pak vzdálenost bodu x_i od středu shluku μ vyjádříme

$$D_i^2 = (x_i - \mu) \Sigma^{-1} (x_i - \mu)^T, \quad (17)$$

kde Σ je kovariační matice setu x .

5 VORONÉHO DIAGRAMY

Voroného diagramy mohou pomoci s geometrickou představou, jak jsou atomy v prostoru uspořádány. Díky tomu si vysloužily místo v této práci.

5.1 Voroného diagram [26]

Je dána množina bodů v prostoru. Voroného diagram je pak model, kde každému bodu přiřadíme jeho nejbližší okolí. Pro každý bod v tomto okolí bude platit, že neexistuje bod z původní množiny, který by mu byl blíže než ten, ke kterému je okolí přiřazeno. Příklad plošného diagramu se nachází na obr. 5-1.

Voroného diagramy vzbudily zájem díky tomu, že popisují velké množství přirozených situací. Lidskému vnímání například pomáhají strukturovat myšlenky a vnímat problém na vyšší úrovni. Někteří autoři dokonce tvrdí, že se jedná o nejzákladnější konstrukci definovanou množinou diskrétních bodů. Další text se bude věnovat zavedení Voroného diagramů pomocí řádných definic.

Je dána množina n bodů zvaná S . Dále jsou dané dvě pozice $p, q \in S$, dominance p nad q je definována jako podmnožina prostoru, která je blíže k p či q . Formálně se zapisuje

$$\text{dom}(p, q) = \{x \in \mathbb{R}^d \mid \delta(p, q) \leq \delta(x, q), \quad (18)$$

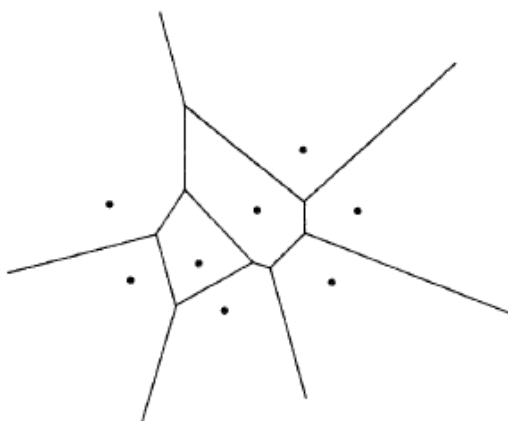
kde d je počet dimenzí a δ je klasická euklidovská vzdálenost.

$$\delta(x, y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2}. \quad (19)$$

Zjevně lze říci, že $\text{dom}(p, q)$ je poloprostor oddělený rovinou, pro kterou platí, že všechny její body jsou stejně vzdáleny od obou bodů p, q .

Oblast náležící bodu p je část prostoru ležící ve všech dominancích p nad ostatními body.

$$\text{obl}(p) = \bigcap_{q \in S - \{p\}} \text{dom}(p, q) \quad (20)$$



Obrázek 5-1 Voroného diagram [26]

Tyto oblasti jsou mnohoúhelníky ve 2D a mnohostěny ve 3D prostoru. Jedná se vždy o konvexní množiny bodů.

Tímto přístupem bude každý bod prostoru přiřazen některé z oblastí nebo její hranici.

Prostor se rozdělí na přesně n oblastí, Některé z nich budou jednoznačně neomezené.

Oblast vypovídá hodně informací o bodu, ke kterému je přiřazena. Lze je vyčíst například z obsahu oblasti či počtu hran (stěn).

5.2 Rovina mezi body

Jedna z prvních věcí, které budou potřeba, je výpočet roviny, která je mezi dvěma body X a Y .

Nejdříve se ověří, zda se nejedná o body se stejnými souřadnicemi, v takovém případě by daný výpočet nedával smysl.

Bude se využívat parametrický zápis roviny ρ , tedy budou potřeba bod S a dva směrové vektory \vec{u}, \vec{v} roviny ρ .

$$\rho = S + k\vec{u} + l\vec{v} \quad k, l \in \mathbb{R} \quad (21)$$

Bod S se zvolí jednoduše v polovině úsečky mezi zkoumanými body.

$$S = \frac{(X + Y)}{2} \quad (22)$$

Směrové vektory roviny jsou kolmé na vektor $\vec{s} = X - Y$, pro jejich skalární součin tedy platí

$$\vec{s} \cdot \vec{u} = s_1 u_1 + s_2 u_2 + s_3 u_3 = 0. \quad (23)$$

V algoritmu je výpočet realizován tak, že první dvě souřadnice jsou zvoleny rovny jedné a poslední se dopočítá.

$$\vec{u} = (1, 1, -\frac{(s_1 + s_2)}{s_3}) \quad (24)$$

Musí se však dát pozor na případ, kdy $s_3 = 0$. Pak by byl zvolen vektor $\vec{u} = (0, 0, 1)$.

Vektor \vec{v} se zvolí tak, aby byl kolmý na vektor \vec{u} , tato vlastnost se pak bude hodit později. K výpočtu se využije vektorový součin.

$$\vec{v} = \vec{s} \times \vec{u} \quad (25)$$

Nakonec se z \vec{u} a \vec{v} udělají jednotkové vektory kvůli dalšímu použití.

5.3 Průnik rovin

Jsou dané dvě parametricky zadané roviny

$$\rho = A_1 + k_1 \vec{u}_1 + l_1 \vec{v}_1 \quad k_1, l_1 \in \mathbb{R}, \quad (26)$$

$$\sigma = A_2 + k_2 \vec{u}_2 + l_2 \vec{v}_2 \quad k_2, l_2 \in \mathbb{R}. \quad (27)$$

Hledá se rovnice přímky p v parametrickém tvaru, která by byla průnikem těchto rovin (obr. 5-2). Její směrový vektor \vec{s} je pak kolmý na normálové vektory rovin (26) a (27). Vektor \vec{s} vypočítáme následovně.

$$\vec{n}_1 = \vec{u}_1 \times \vec{v}_1 \quad (28)$$

$$\vec{n}_2 = \vec{u}_2 \times \vec{v}_2 \quad (29)$$

$$\vec{s} = \vec{n}_1 \times \vec{n}_2 \quad (30)$$

Pak už zbývá pouze najít některý z bodů společných pro roviny ρ, σ , pro který platí rovnice

$$A_1 + k_1 \vec{u}_1 + l_1 \vec{v}_1 = A_2 + k_2 \vec{u}_2 + l_2 \vec{v}_2 \quad k_1, l_1, k_2, l_2 \in \mathbb{R}. \quad (31)$$

Z tohoto vztahu se získá soustava tří rovnic o čtyřech neznámých, což by pro tento případ bylo zbytečně nejednoznačné a komplikované. Vytvoří se proto rovnice pro již jednoznačně určený bod.

Tento bod bude průnikem roviny ρ a přímky q náležící rovině σ kolmé na vektor \vec{s} .

$$q = A_2 + k (\vec{s} \times \vec{n}_2) \quad k \in \mathbb{R} \quad (32)$$

Potom je možné používat zjednodušenou rovnici

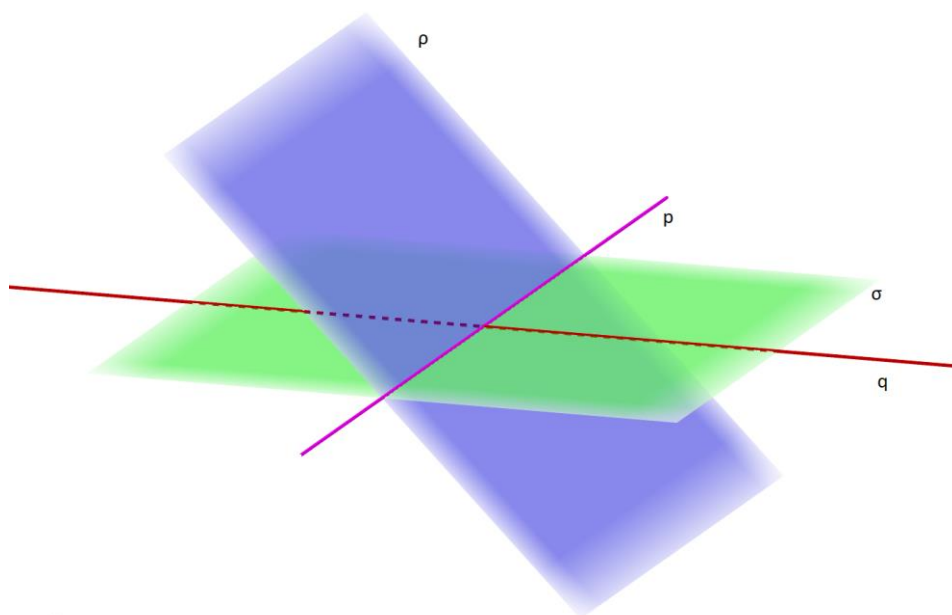
$$A_1 + k_1 \vec{u}_1 + l_1 \vec{v}_1 = A_2 + k (\vec{s} \times \vec{n}_2), \quad k_1, l_1, k \in \mathbb{R} \quad (33)$$

ze které vypočteme bod A .

$$A = A_2 + k (\vec{s} \times \vec{n}_2), \quad k \in \mathbb{R} \quad (34)$$

Rovnice hledané přímky je pak

$$p = A + k \vec{s} \quad k \in \mathbb{R}. \quad (35)$$



Obrázek 5-2 Průnik rovin

5.4 Transformace souřadnic

V průběhu algoritmu se bude pro usnadnění transformovat souřadnice bodů a vektorů z trojrozměrných na dvourozměrné.

5.4.1 Bod

Je zadáný bod X a rovina

$$\rho = A + k\vec{u} + l\vec{v} \quad k, l \in \mathbb{R}. \quad (36)$$

Bod X leží v rovině ρ .

Stanoví se nová soustava souřadnic s počátkem v bodě A a osami ve směru vektorů \vec{u} a \vec{v} . V tomto momentě je velkou výhodou, že všechny roviny jsou určeny ortogonálními směrovými vektory.

Úkolem je najít nové souřadnice bodu X ve stanovené soustavě souřadnic. Jelikož platí $|\vec{u}| = |\vec{v}| = 1$, lze říci, že k, l z rovnice

$$A + k\vec{u} + l\vec{v} = X, \quad (37)$$

jsou transformovanými souřadnicemi bodu X bez zkreslení velikostí.

Pro lepší přiblížení situace lze přirovnat tuto soustavu k běžné kartézské. Místo bodu A se použije běžný počátek $(0, 0)$ a místo \vec{u}, \vec{v} pak vektory $(1, 0), (0, 1)$. Pak si lze představit, že souřadnice bodů jsou určeny tím, kolikrát tyto vektory přičteme k počátku, podobně, jak je naznačeno v rovnici výše.

5.4.2 Vektor

Je zadáný vektor \vec{x} a rovina

$$\rho = A + k\vec{u} + l\vec{v} \quad k, l \in \mathbb{R}. \quad (38)$$

Vektor \vec{x} je směrnice vektor roviny ρ .

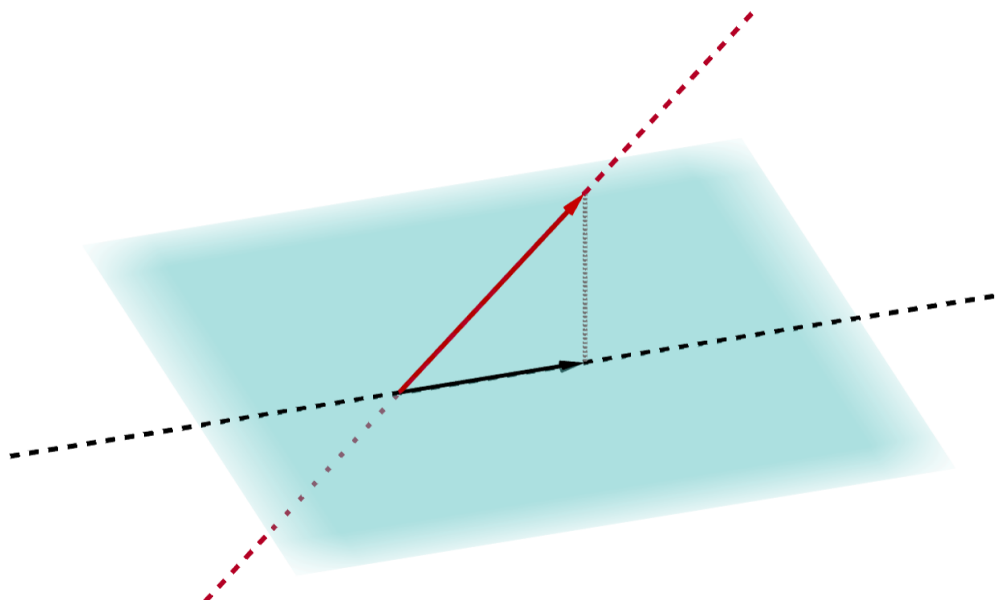
Postupuje se velmi podobně, jako v případě s bodem. Jediný rozdíl je v tom, že vzhledem k tomu, že vektor je určen pouze směrem, nikoliv svým počátkem, určí se jeho souřadnice pouze jako kombinace vektorů bez ohledu na počátek A. Rovnice má pak tvar

$$k\vec{u} + l\vec{v} = \vec{x}, \quad (39)$$

kde k, l jsou transformovanými souřadnicemi vektoru \vec{x} .

5.5 Zobrazení vektoru na rovinu

Tato kapitola se zabývá problémem, jak zobrazit vektor na plochu, přestože k ní není směřový. Situaci je možné si představit například jako promítání, kdy se nad vektorem rozsvítí světlo a jeho zobrazení na rovinu představuje vržený stín. (obr. 5-3)



Obrázek 5-3 Zobrazení vektoru na rovinu

V tomto případě bude stačit pravoúhlé promítání.

Je dána rovina

$$\rho = A + k\vec{u} + l\vec{v} \quad (40)$$

a vektor \vec{s} . Normálový vektor roviny se vypočítá vektorovým součinem.

$$\vec{n} = \vec{u} \times \vec{v} \quad (41)$$

Vektor \vec{s} je pak lineární kombinací ostatních vektorů.

$$k\vec{u} + l\vec{v} + m\vec{n} = \vec{s} \quad (42)$$

Po vyřešení této rovnice je možné vypočítat souřadnice promítnutého vektoru na plochu ρ

$$\vec{s}_{\text{prom}} = k\vec{u} + l\vec{v} \quad (43)$$

nebo rovnou získat jeho transformované souřadnice

$$\vec{s}_{\text{transf}} = (k, l). \quad (44)$$

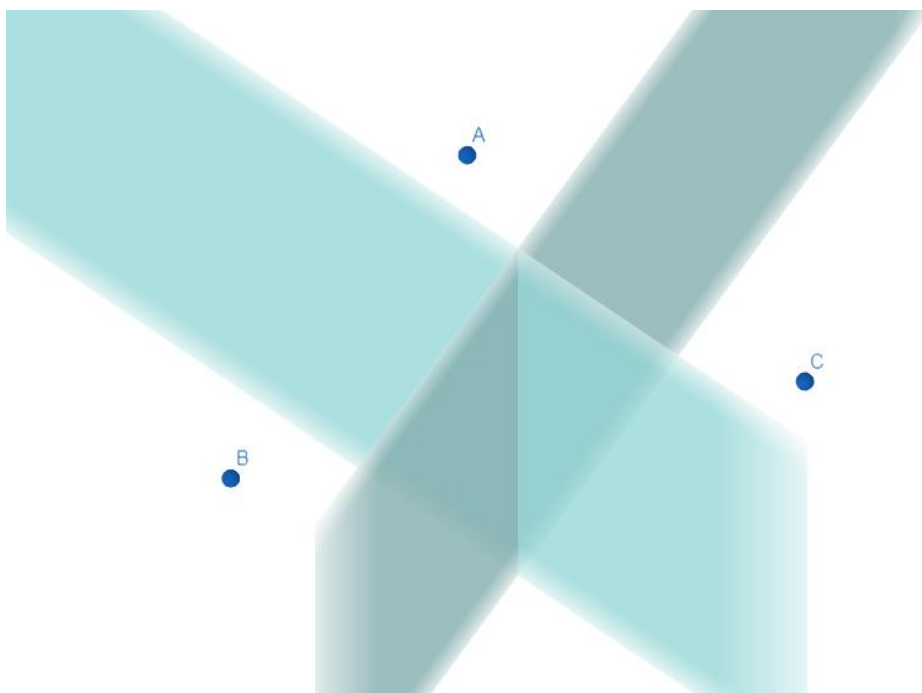
5.6 Algoritmus

Následující algoritmus nám pomůže najít plochu mezi dvěma body soustavy, které budeme dále označovat jako A, B.

Po zvolení bodů určíme rovnici roviny, která rozděluje prostor na dva poloprostory. Pro ty platí, že všechny jejich body jsou blíže buď bodu A nebo bodu B, podle toho, kterému z nich je poloprostor přiřazen. Rovina bude dále označována ρ .

Na této rovině se bude ohraničovat útvar, který je blíže bodům A, B oproti ostatním bodům soustavy. Hranice budou tvořeny přímkami.

K bodům A, B se v průběhu algoritmu přidávají další body ze soustavy. Pro přímky pak platí, že jsou průnikem roviny σ a roviny mezi bodem A (případně B, v tomto případě bychom neměli zaznamenat rozdíl) a nově přidaným bodem, který se bude nazývat C.



Obrázek 5-4 Hraniční roviny

Po přidání nového bodu se nejdříve vyšetří, zda bod neleží na přímce AB, v tom případě by pro něj platila rovnice

$$C = A + k \overrightarrow{AB} \quad k \in \mathbb{R}. \quad (45)$$

Pokud ano, rozeznávají se dvě možnosti. Pokud bod C leží mezi body A, B, tedy platí

$$k \in (0, 1). \quad (46)$$

Pak je jasné, že útvar s požadovanými vlastnostmi nemůže existovat, protože by jeho body byly vždy blíže bodu C než bodu A nebo B.

Pokud bod leží mimo úsečku AB, nijak nepřispívá k vytvoření omezení útvaru, protože jeden z původních bodů bude na celé rovině ρ zaručeně blíže. V tomto případě je možné okamžitě přejít k přidání nového bodu.

Mimo tento specifický případ se určí rovina mezi body A, C a zjistí se její průnik s rovinou ρ (obr. 5-4), čímž se získá rovnice přímky p .

$$p = A_p + k\vec{u}_p \quad k \in \mathbb{R}. \quad (47)$$

Rovnice však není vše, co je o této přímce potřeba vědět. Důležité je také to, jakým směrem od přímky se nachází bod A, což určíme vektorem

$$\vec{s} = A - C. \quad (48)$$

Poslední důležitou informací je omezení přímky, protože hranici útvaru bude tvořit pouze úsečka ležící na této přímce. Zaznamenáme jej jako interval pro k (z (45)), který pro začátek bude

$$k \in (-\infty, \infty), \quad (49)$$

v průběhu algoritmu se budou hranice měnit.

Na rovnici přímky se pak aplikuje transformace souřadnic, aby se dostaly pouze dvourozměrné souřadnice odpovídající soustavě, ve které osy mají směr směrových vektorů roviny ρ . Tyto vektory jsou ortogonální a jednotkové, proto lze tuto úpravu udělat jednoduše a pokračovat v dalších výpočtech efektivněji. Vektor určující, jakým směrem od přímky se nachází žádaný poloprostor, se zobrazí na rovinu ρ .

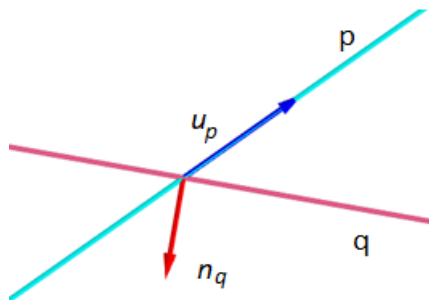
Pokud je p první přidaná přímka, je možné pokročit k přidání nového bodu. Pokud však již na seznamu nějaké přímky jsou, zjišťuje se, jak se přímky navzájem omezují.

Ze seznamu se vezme jedna přímka po druhé a porovnává se s novou přímkou p . Nejříve se najde průnik přímek

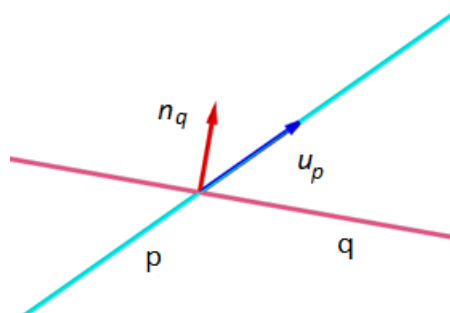
$$A_p + k\vec{u}_p = A_q + l\vec{u}_q, \quad (50)$$

kde q je aktuálně zkoumaná přímka ze seznamu. Zaznamenáme konstanty k, l .

Nyní se musí zjistit, jakým způsobem ovlivní tento průnik hranice přímek. Všeobecně lze říci, že pokud je úhel mezi vektory \vec{u}_p a \vec{n}_q menší než pravý úhel, zkoumá se minimum pro k , a pokud je větší, zkoumá se maximum.



Obrázek 5-5 Příklad $\cos(\alpha) < 0$

Obrázek 5-6 Příklad $\cos(\alpha) > 0$

Úhel lze určit pomocí vzorce

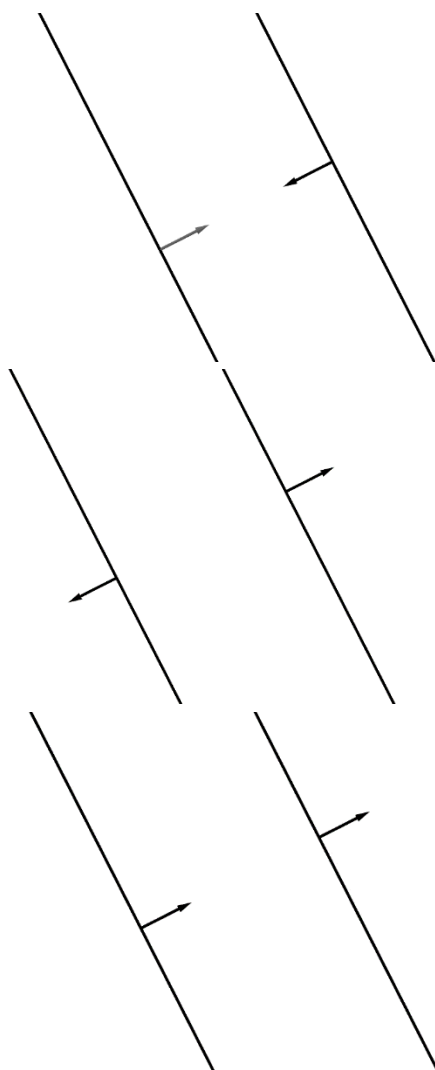
$$\cos(\alpha) = \frac{\vec{u}_p \cdot \vec{n}_q}{|\vec{u}_p| |\vec{n}_q|} \quad (51)$$

Podstatné je pouze znaménko daného výrazu, lze proto výpočet zjednodušit na

$$c = \cos(\alpha) = \vec{u}_p \cdot \vec{n}_q. \quad (52)$$

Dále se musí dát pozor na situaci, kdy vyjde $c < 0$ (úhel $\alpha > 90^\circ$, zkoumá se maximum, obr. 5-5) a k je menší než aktuální minimum funkce. V takovém případě platí, že přímka p není hranicí, a proto ji nebudeme přidávat do seznamu hranic množiny. Pokud tento případ nenastane, porovnájí se se současným maximem a případně jej změníme. Podobně aplikujeme pro $c > 0$ (obr. 5-6).

Speciální případ pak nastává pro $c = 0$, v tomto případě jsou přímky rovnoběžné. Zde existuje více možností, jak situace vypadá (obr. 5-7). Pokud vektory \vec{s} mají opačný směr, ale směřují každý od přímky, ke které náleží, do prostoru mezi přímkami, lze obě přímky ponechat bez úpravy omezení. Pokud směřují vně, znamená to, že průnik polorovin je nulový a mezi body A, B hledaný útvar neexistuje. Pokud směrové vektory mají stejný směr, znamená to, že se ponechá pouze přímku bližší bodu A.



Obrázek 5-7 Rovnoběžné vektory

Pseudokód by pak vypadal následovně:

```
c = up . sq
IF (c > 0 AND k < maxp)
  minp = MAX(minp, k)
  PŘIDAT p
ELSE IF (c < 0 AND k > maxp)
  maxp = MIN(maxp, k)
  PŘIDAT p
ELSE IF (c = 0)
  NEPŘIDÁVAT p
ELSE
  IF (sp[1] * sq[1] > 0)
    IF (|A, p| < |A, q|)
      MAZAT q
    ELSE
      NEPŘIDÁVAT p
  ELSE
    IF (|A, p| + |A, q| = |p, q|)
      VRÁTIT PRÁZDNÝ LIST
    ELSE
      PŘIDAT p
```


6 ANALÝZA

6.1 Popis dat

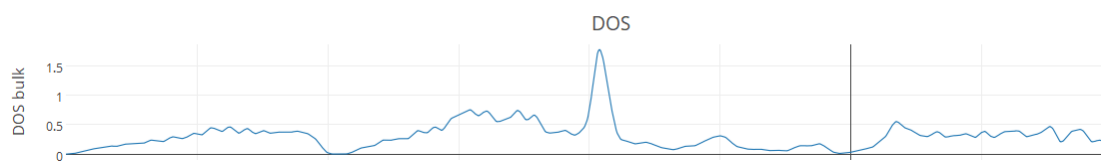
Data k analýze se skládají z šesti csv souborů, tři pro křivky DOS a tři pro křivky COHP. První z těchto souborů obsahuje data k primitivní buňce krystalové mřížky, druhý ke stabilní struktuře hranice zrn a poslední k nestabilní.

Tabulky v souborech obsahují souřadnice jednotlivých bodů křivek. Jako osa x zde se může označit sloupec Energie. Tyto hodnoty jsou pro všechny křivky společné a pravidelně rozmístěné. Osa y pak odpovídá datům daných křivek.

Křivku DOS lze přiřadit k určitým atomům, zatímco COHP křivky vyjadřují vztah mezi dvěma atomy.

6.2 Analýza DOS křivek

Na obr. 6-1 a 6-2 je vidět DOS křivky pro primitivní buňku, které slouží u některých analýz jako vzor k porovnání.



Obrázek 6-1 DOS pro Si primitivní buňky



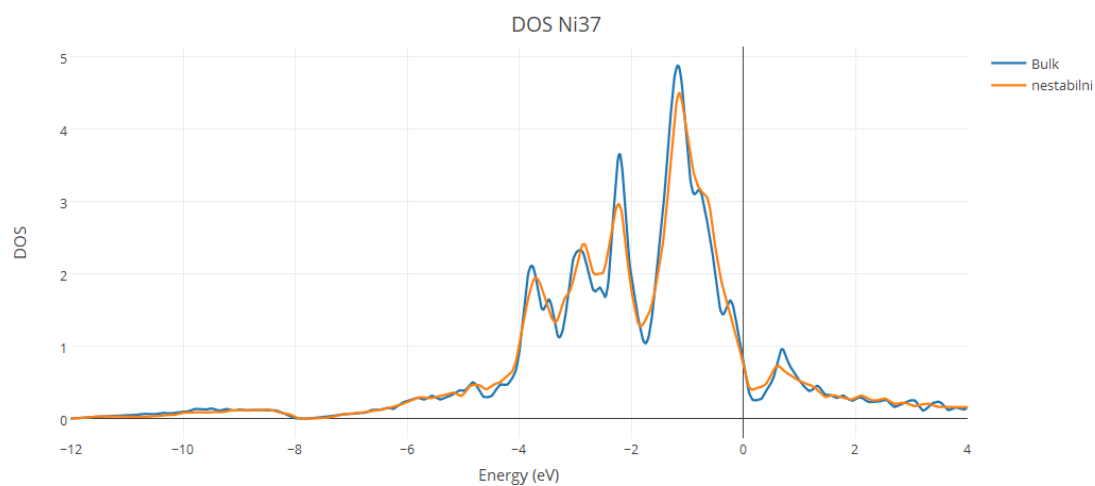
Obrázek 6-2 DOS pro Ni primitivní buňky

6.2.1 Korelace

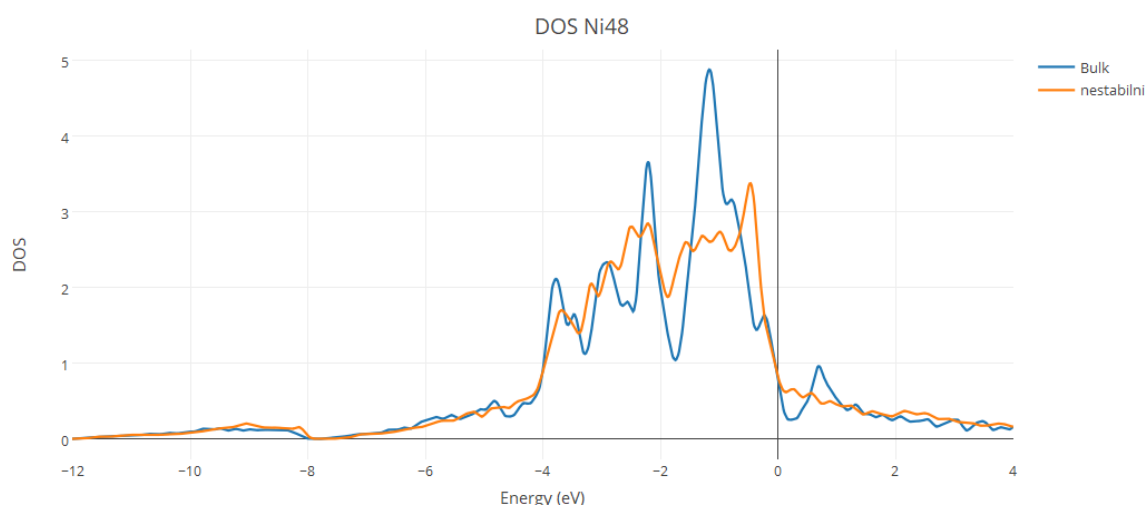
Jedním ze základních kroků při zkoumání datasetu je vypočtení korelací. Hlavní otázkou ovšem je, mezi kterými křivkami korelaci vyhledávat. Jelikož všechny atomy stabilní varianty nemusí mít odraz ve variantě nestabilní, budou se porovnávat s atomy primitivní buňky krystalické mřížky. Není snad nutno dodávat, že se porovnávají atomy Ni s Ni a Si s Si.

Všechny korelace vyšly poměrně vysoké (nejnižší korelace byla 0,832), přesto malý rozdíl stačí pro rozdělení křivek.

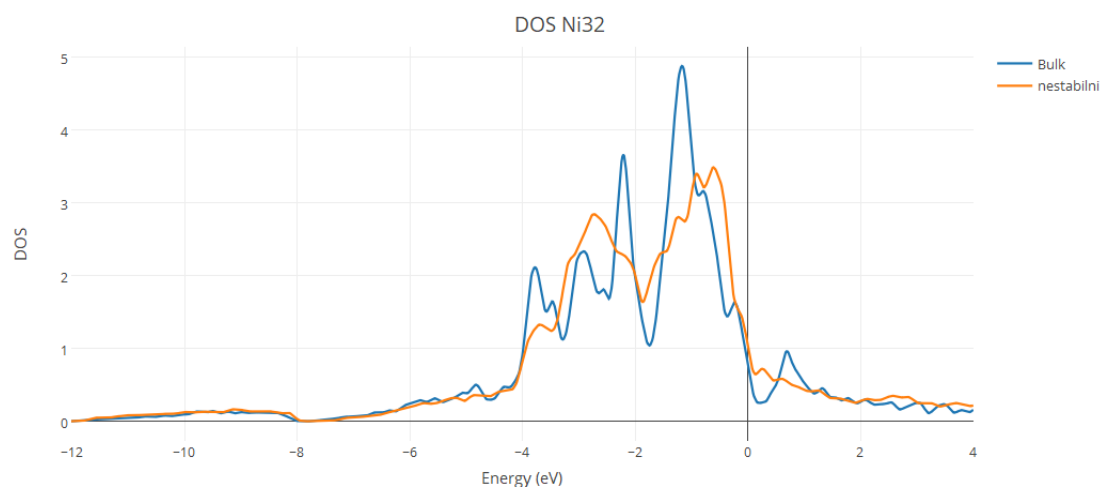
Vybrané křivky, u kterých vyšla korelace vysoká, jsou vykresleny na obr. 6-3 a křivky s nižší korelací na obr. 6-4 a obr. 6-5. Již od pohledu lze říci, že rozdíl je zde zřejmý a tuto metodu lze označit za vhodnou pro analýzu křivek.



Obrázek 6-3 DOS pro Ni37, korelace = 0,988



Obrázek 6-4 DOS pro Ni48, korelace = 0,91



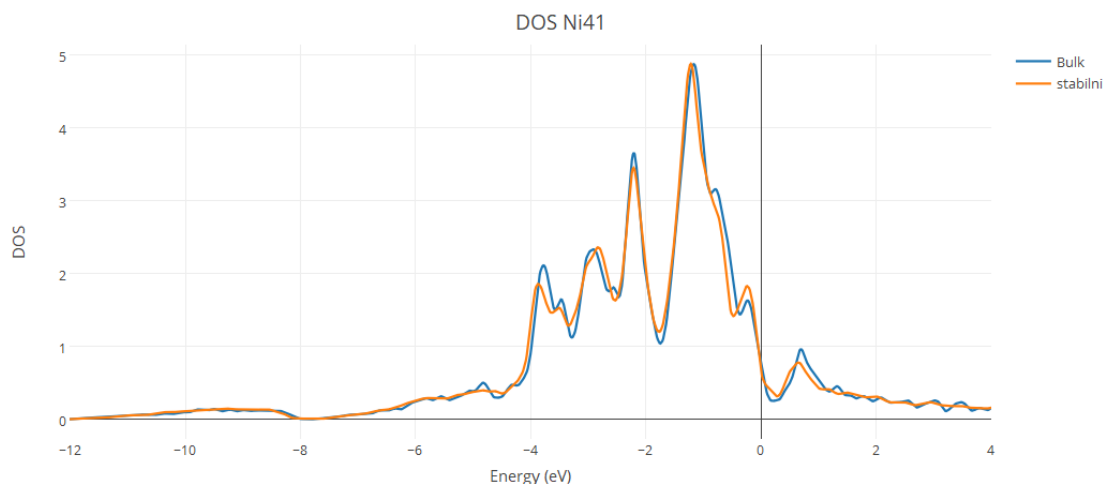
Obrázek 6-5 DOS pro Ni32, korelace = 0,911

6.2.2 Vzdálenost křivek

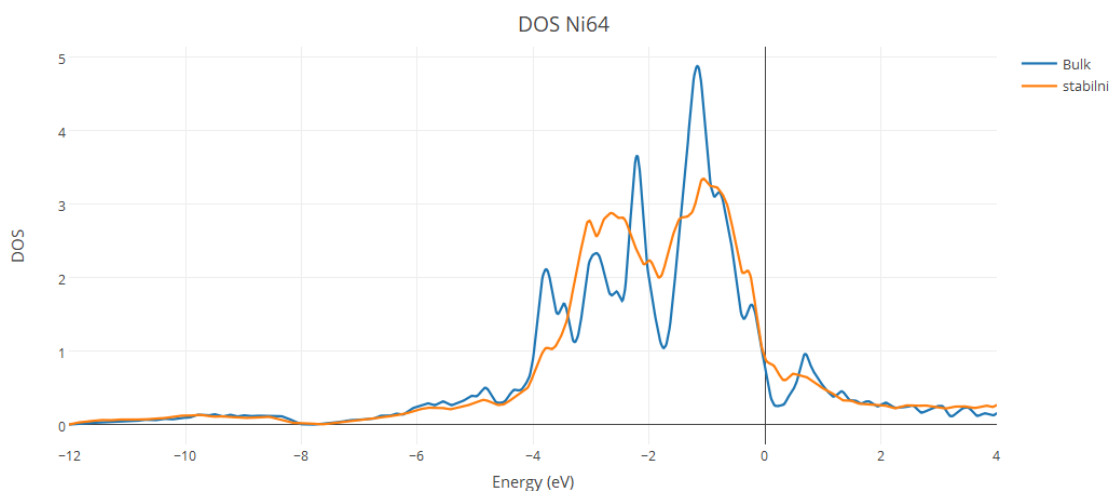
Další způsob, kterým lze porovnat křivky s těmi z primitivní mřížky, je určit vzdálenost mezi nimi. Vzdálenost se bude měřit pouze ve směru y-ové osy.

Nejjednodušší možnost je spočítat vzdálenost křivek pro každou zadanou energii a poté z nich udělat průměr.

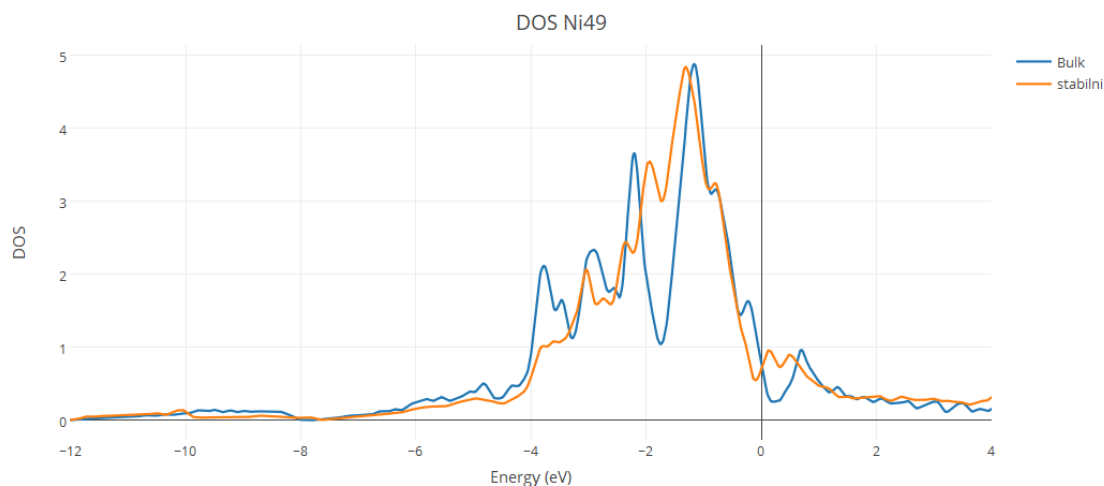
I tato metoda se ukázala být vcelku úspěšná. Na obr. 6-6 je zobrazena křivka s nízkou vzdáleností od svého vzoru a je viditelné, že i vizuálně si jsou křivky poměrně podobné. Případ větší vzdálenosti je možné zhodnotit na obr. 6-7 a obr. 6-8.



Obrázek 6-6 DOS pro Ni41, vzdálenost = 0,07



Obrázek 6-7 DOS pro Ni64, vzdálenost = 0,2



Obrázek 6-8 DOS pro Ni49, vzdálenost = 0,22

Mějme hypotézu, že čím blíže je energie k nule, tím stoupá důležitost přiřazené hodnoty. Je možné tedy udělat vážený průměr s lineární klesající váhou vzhledem ke vzdálenosti od nuly.

$$\text{Váha}_i = \frac{\max(|E|) - |E_i|}{\max(|E|)}$$

V našem případě nebyla zaznamenána významná změna mezi váženým a běžným průměrem, proto se dále využívá pouze průměr nevážený.

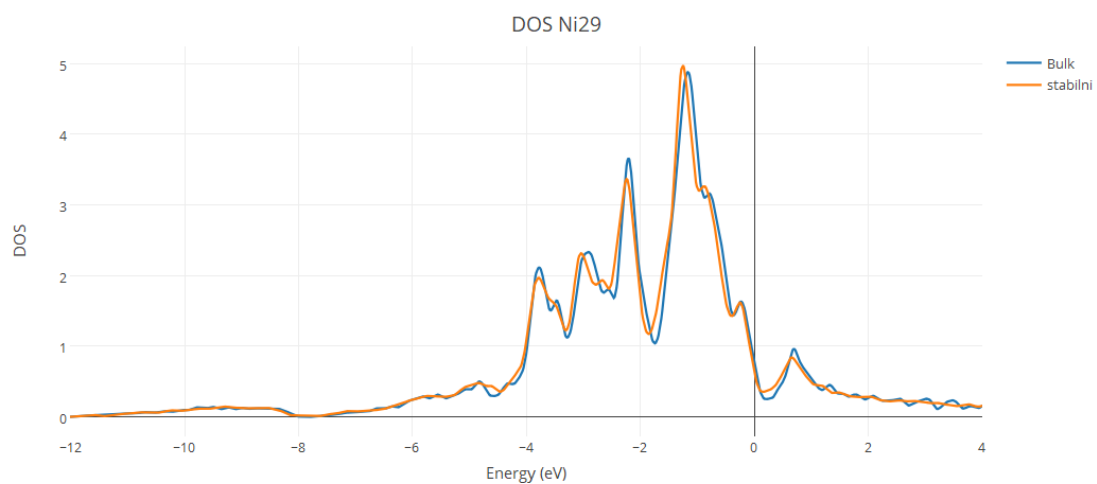
6.2.3 Vzdálenost pomocí DTW

U některých křivek je možné si všimnout, že vrcholy, které by si případně mohly odpovídat nejsou zarovnané pod sebe, jak je vidět na obr. 6-8.

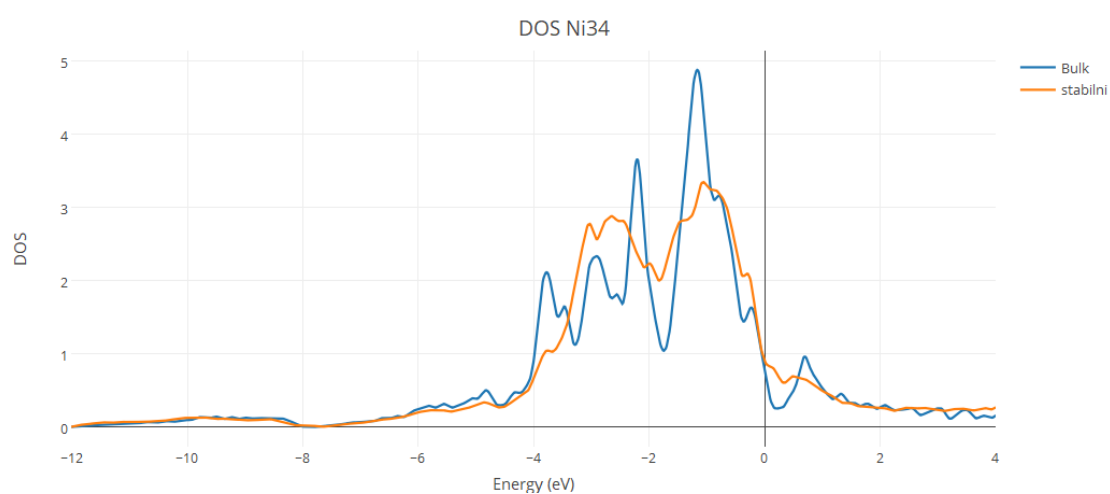
Pokud je podstatné odlišit případ na obr. 6-7, kde je evidentní, že křivky si nejsou nikterak podobné, a případ na obr. 6-8, kde se jedná o podobné křivky, které jsou jen o kousek posunuté v rámci osy x, je zřejmé, že se bude muset využít jiné metody.

Z toho důvodu se aplikuje na dataset nejdříve metoda DTW, ze které se dostane dvojice bodů, které spolu korespondují. Jako vzdálenost křivek se určí průměrná vzdálenost odpovídajících si bodů.

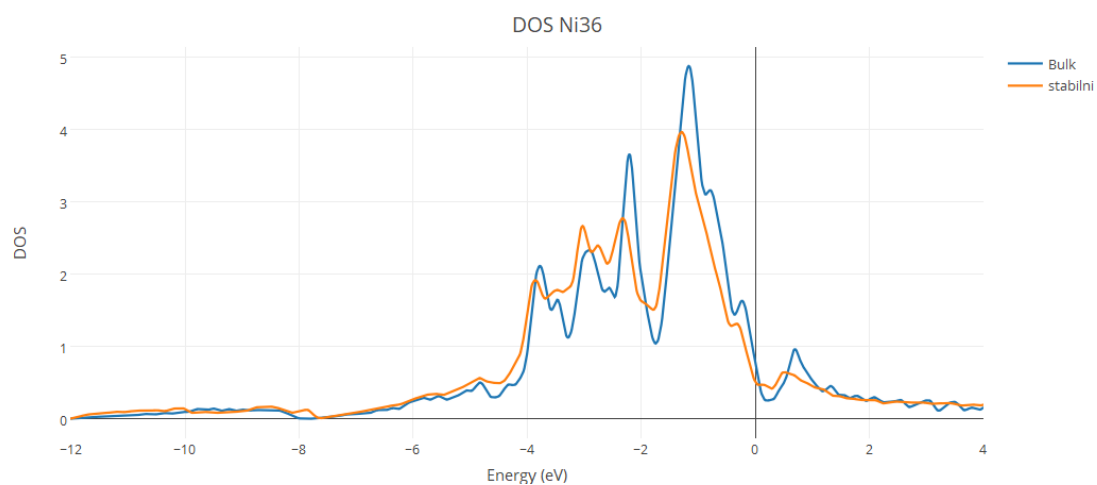
Na obr. 6-9 je možné vidět ideální případ, kde je vzájemná vzdálenost nízká. Na obr. 6-10 je varianta s nejvyšší vzdáleností. Na obr. 6-11 je křivka posunutá a můžeme vidět, že pomocí této metody se výrazně redukovala vzdálenost oproti výsledkům v kapitole 6.2.2.



Obrázek 6-9 DOS pro Ni29, vzdálenost = 0,013



Obrázek 6-10 DOS pro Ni34, vzdálenost = 0,06



Obrázek 6-11 DOS pro Ni36, vzdálenost = 0,042

6.2.4 Detekce odlehklých křivek pomocí horizontálních vrstevnic

Poslední použitá metoda má výhodu v tom, že nepotřebuje vzor v podobě křivky ze základní mřížky. Pomocí Detekce odlehklých křivek pomocí horizontálních souřadnic se získá hodnota excentricity dané křivky oproti ostatním v tomtéž datasetu.

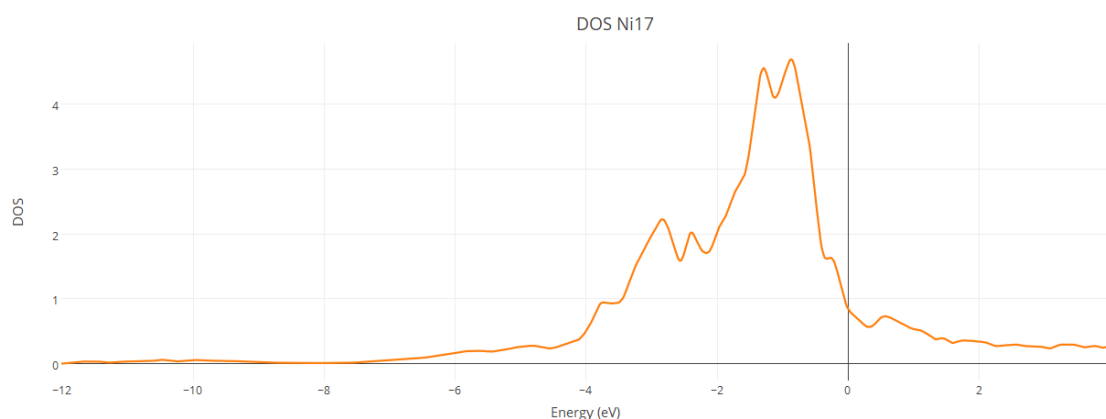
Zároveň tato metoda neporovnává vzdálenost křivek, ale čistě jejich tvar. Jeho povaha je tedy odlišná od výše popsanych metod.

Pro aplikaci metody je samozřejmě nutné rozdělit na křivky náležící Ni a ty náležící Si. Po vícero pokusech byla zvolena hodnota pro počet horizontálních vrstevnic rovna $n = 10$.

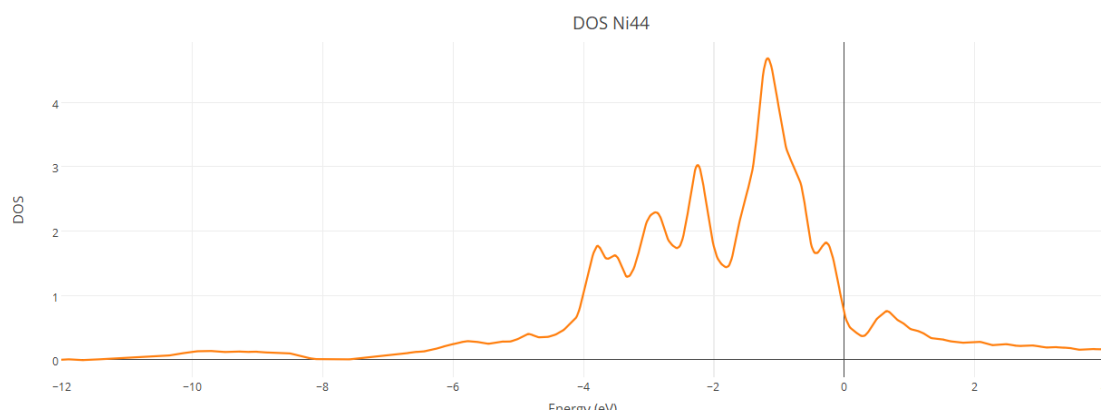
Jako nejvíce excentrická byla určena křivka na obr. 6-12, jako nejprůměrnější pak křivka na obr. 6-13. Lze si všimnout, že křivka na obr. 6-13 je velmi podobná té z primitivní mřížky.

Křivka Ni17 (obr. 6-12) byla i předchozími metodami vyhodnocena jako více či méně vyčnívající. Dá se tedy říci, že potvrzuje výsledek dvou předchozích analýz.

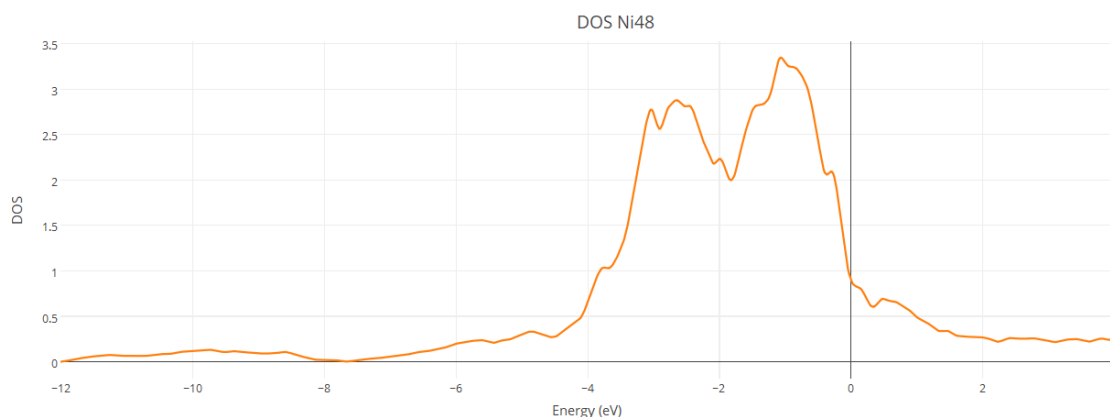
Křivka Ni48 (obr. 6-14) se od ostatních odlišuje tím, že ve všech předchozích metodách poměrně vystupovala, ale u Detekce odlehklých křivek pomocí horizontálních vrstevnic získala pouze průměrné skóre.



Obrázek 6-12 DOS pro Ni17, excentricita = 4,52



Obrázek 6-13 DOS pro Ni44, excentricita = 0,22

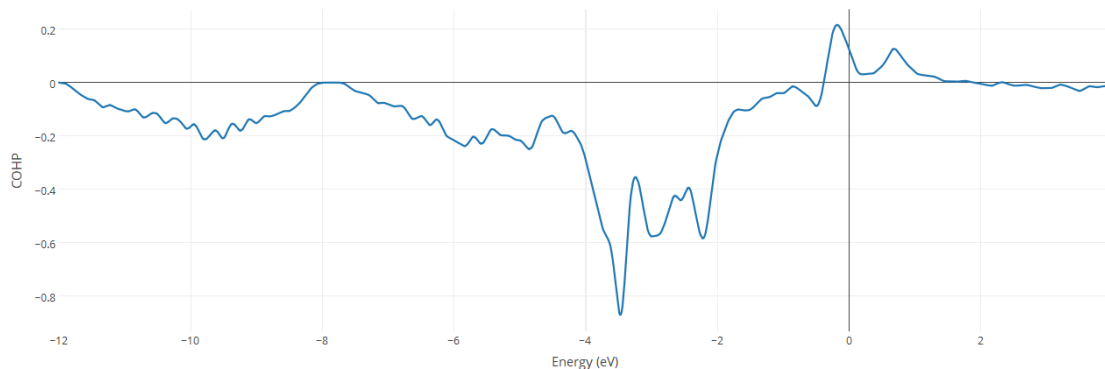


Obrázek 6-14 DOS pro Ni48, excentricita = 2,54

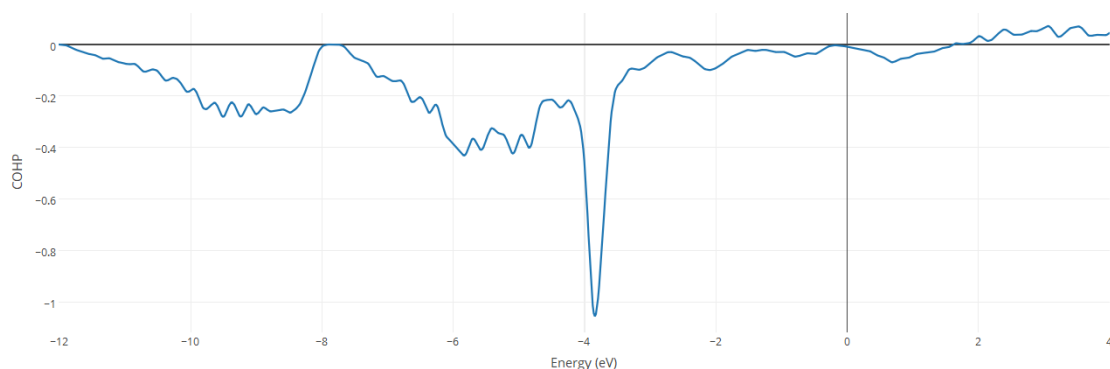
6.3 Analýza COHP křivek

Další text se bude zabývat analýzou a porovnáváním křivek COHP.

Na obr. 6-15 a 6-16 si lze prohlédnout COHP křivky pro primitivní buňku, které slouží u některých analýz k porovnání.



Obrázek 6-15 COHP pro Ni-Ni z primitivní mřížky



Obrázek 6-16 COHP pro Si-Ni z primitivní mřížky

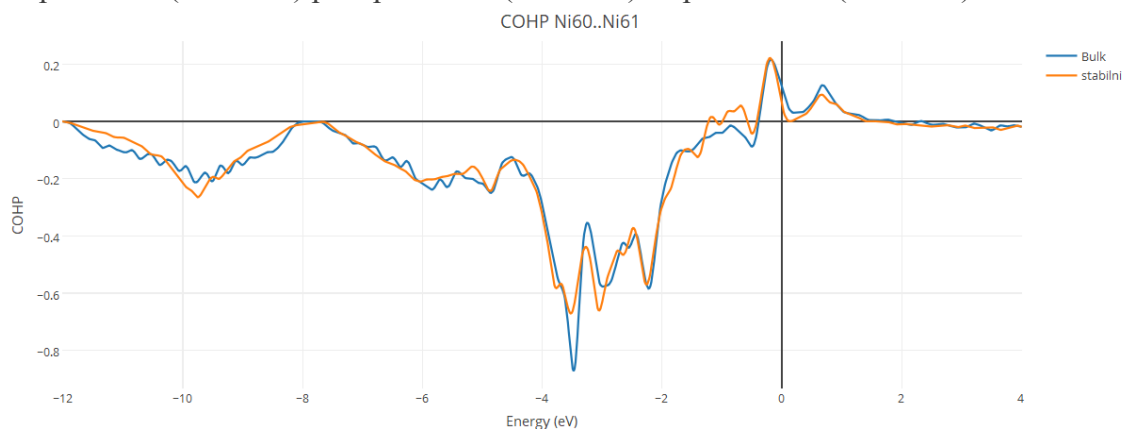
Metody jsou stejné jako v kapitole 6.2, nebude k nim proto uveden úvod, ale rovnou zhodnocení výsledků.

6.3.1 Korelace

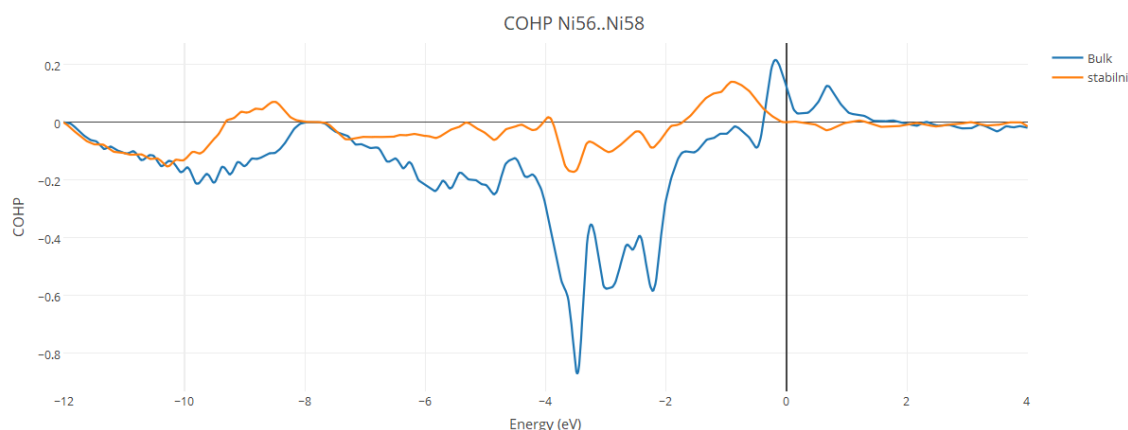
U těchto křivek jsou v korelacích výraznější rozdíly než v kapitole 6.2.1, jejich rozpětí je přibližně $(-0,04; 0,99)$.

Objevují se zde i záporné korelace, ale vzhledem k tomu, jak malou velikost mají, nemusí se jim připisovat významné vlastnosti oproti nízkým kladným. Lze je tedy zjednodušeně brát jako nulové.

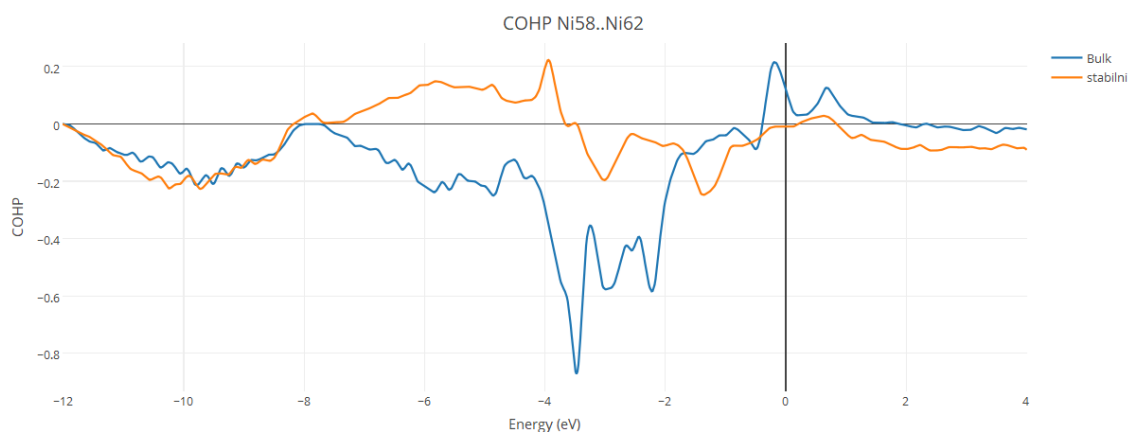
Na grafech je vidět, jak se mění tvar křivky pro různé korelace od téměř stoprocentní (obr. 6.17) přes poloviční (obr. 6.18) až po nulovou (obr. 6.19).



Obrázek 6-17 COHP pro Ni60 – Ni64, korelace = 0,98



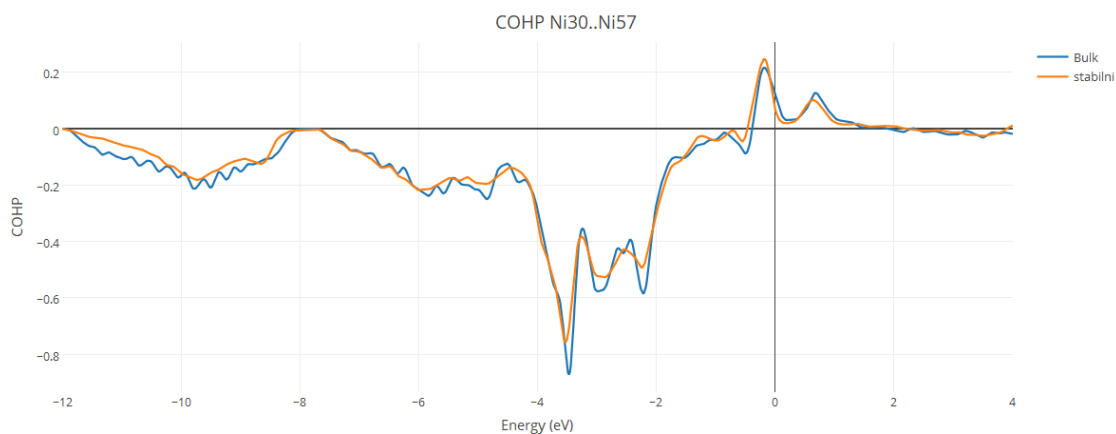
Obrázek 6-18 COHP pro Ni56 – Ni58, korelace = 0,5



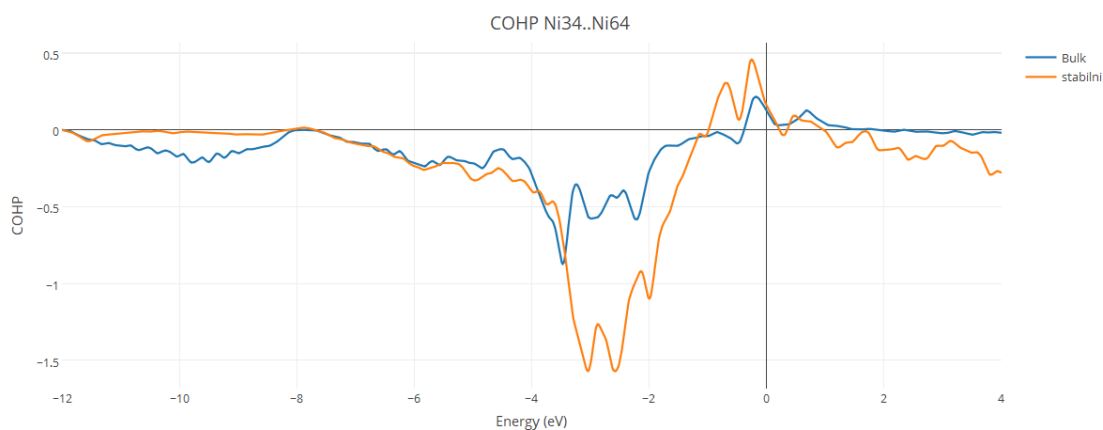
Obrázek 6-19 COHP pro Ni58 – Ni62, korelace = -0,04

6.3.2 Vzdálenost křivek

I u těchto křivek má metoda dobrou vypovídací hodnotu. Zhodnotit to lze podle obrázků 6-20, kde je vzdálenost velmi nízká, a 6-21, kde je vzdálenost poměrně vysoká.



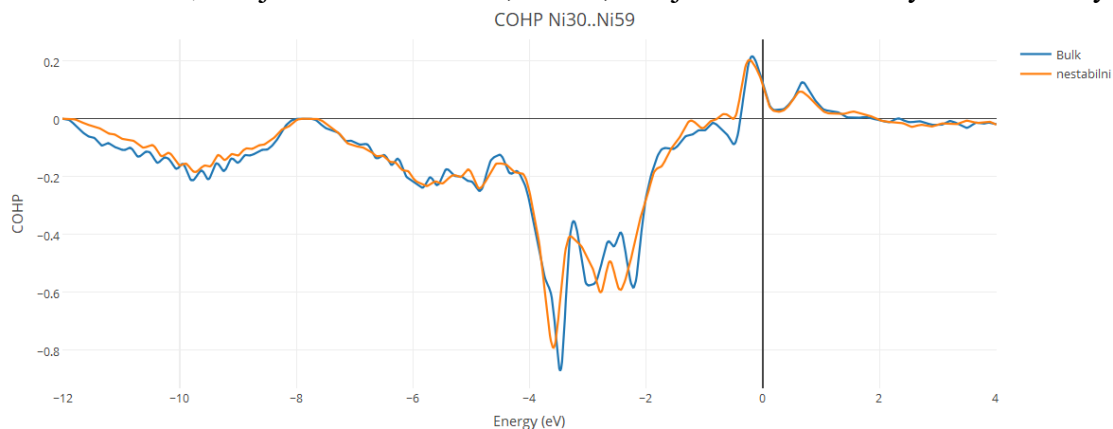
Obrázek 6-20 COHP pro Ni30 – Ni57, vzdálenost = 0,02



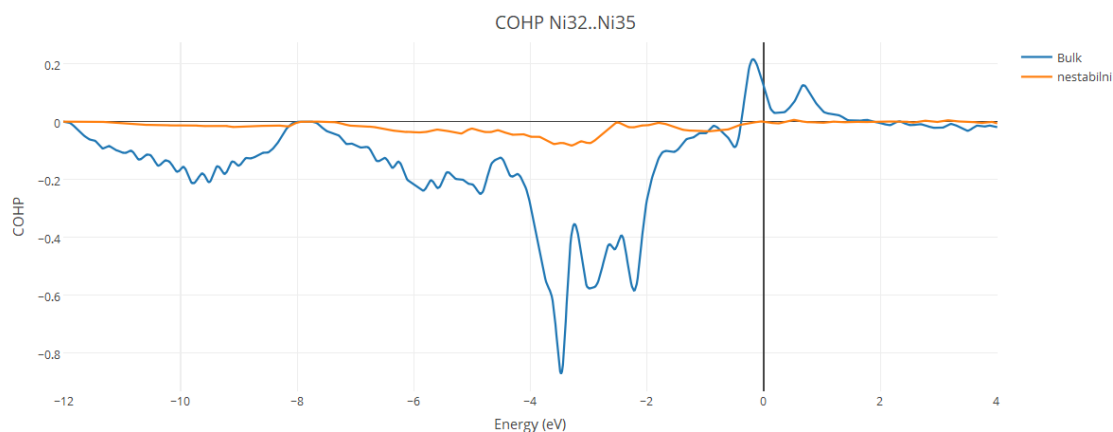
Obrázek 6-21 COHP pro Ni34 – Ni64, vzdálenost = 0,17

6.3.3 Vzdálenost pomocí DTW

Tato metoda zafungovala velmi dobře pro křivky COHP Ni-Ni, jak lze posoudit z obrázků 6-22, kde je vzdálenost nízká, a 6-23, kde jsou od sebe křivky více vzdáleny.



Obrázek 6-22 COHP pro Ni30 – Ni50, vzdálenost = 0,005



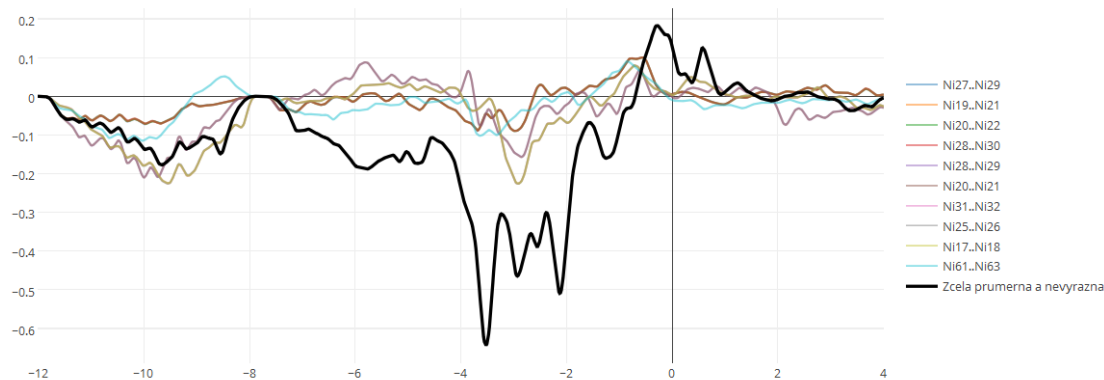
Obrázek 6-23 COHP pro Ni32 – Ni35, vzdálenost = 0,053

6.3.4 Detekce odlehlých křivek pomocí horizontálních vrstevnic

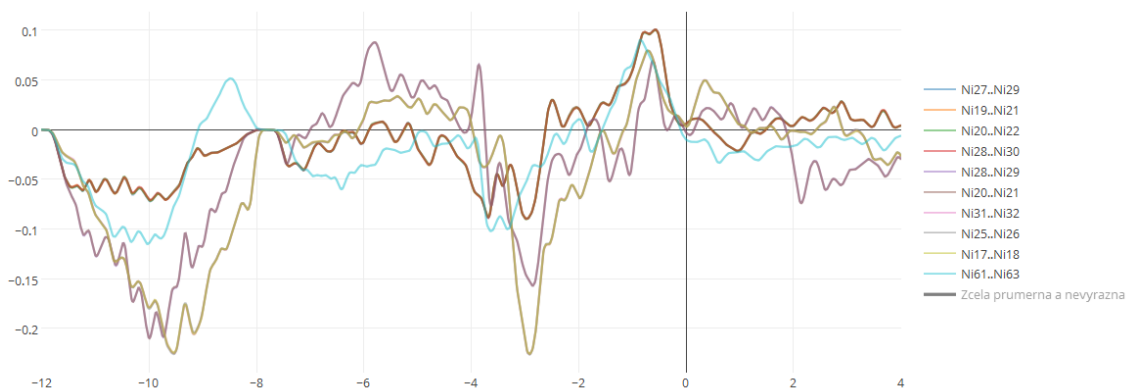
I v tomto případě byla metoda úspěšná. Jako příklad je možné uvést analýzu Ni-Ni z nestabilního souboru.

Na obr. 6-24 je zobrazeno deset nejexcentričtějších křivek a výrazně černě křivka, která byla vyhodnocena jako nejprůměrnější. Pro lepší srovnání tvaru lze nahlédnout na obr. 6-25, kde jsou odlehlé křivky zobrazeny bez té průměrné, aby se dal jednodušeji zhodnotit jejich tvar.

Průměrná křivka je opět velmi podobná té ze základní mřížky. Na první pohled lze říci, že křivky identifikované touto metodou mají skutečně výrazně jiný tvar.



Obrázek 6-24 Odlehlé křivky Ni – Ni, nestabilní



Obrázek 6-25 Odlehlé křivky Ni – Ni, nestabilní, bez průměrné křivky

6.4 Voroného diagramy

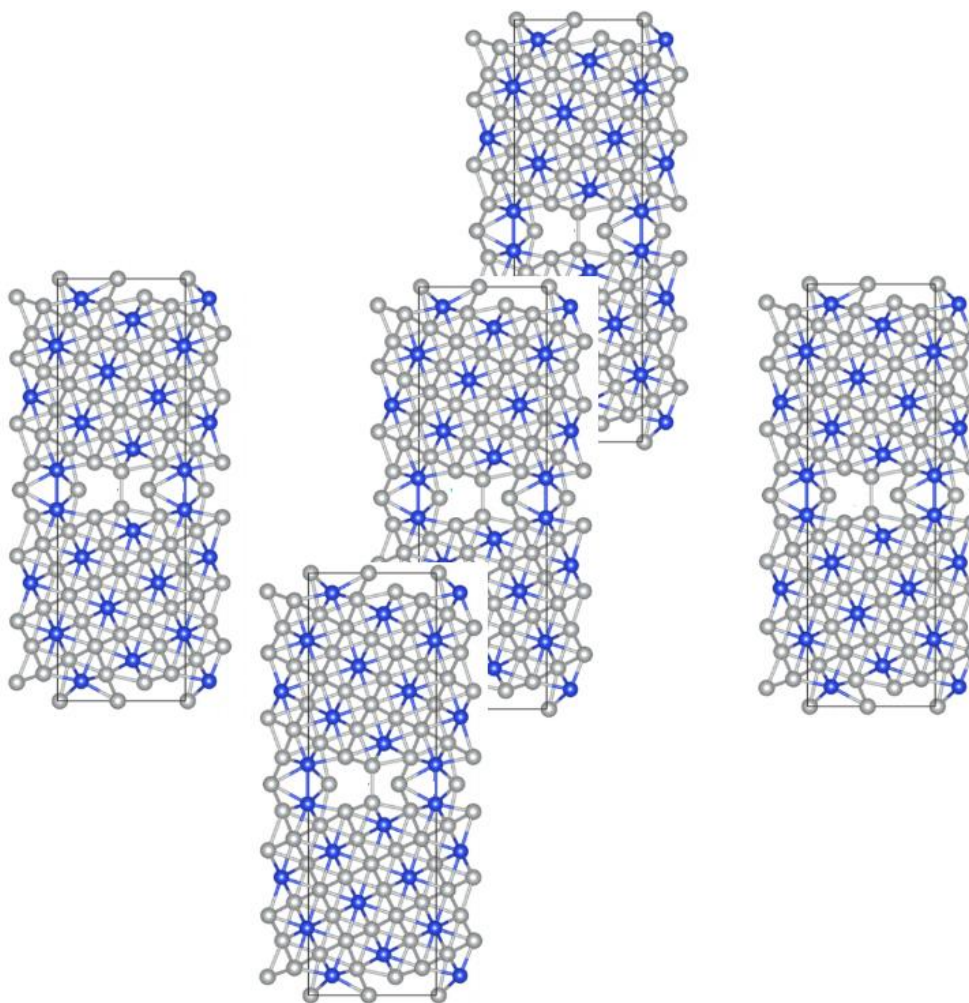
Zajímavé informace lze získat nejen z křivek, ale i z geometrického uspořádání atomů v mřížce. K vyšetření okolí atomu je v této práci použit Voroného diagram (kapitola 5).

Existuje několik možných způsobů, kterými lze z vytvořeného diagramu získat informace, jako například objem útvaru náležící k atomu, jeho obsah nebo tvar. V této práci se využívá obsah stěny útvaru, ten lze přidělit dvěma atomům, mezi kterými tvoří příslušná plocha hranici.

Pokud je tato plocha větší, zpravidla to pak znamená, že jsou atomy blízkými sousedy. Pokud je plocha menší nebo zcela nulová, znamená to, že tyto atomy jsou vzdáleny a v okolí jejich spojnice se nachází atomy, které výrazně ovlivňují (oslabují) jejich interakci.

Pracuje se zde se souborem obsahujícím souřadnice atomů mřížky. Tato mřížka je pouze „základní stavební jednotka“ mřížky větší, která vznikne opakováním té základní.

Aby byl výpočet Voroného diagramu plnohodnotný i pro okraje základní mřížky, musí se do souboru přidat i souřadnice okolních mřížek, jak je naznačeno na obr. 6-26.



Obrázek 6-26 Opakování mřížky [17]

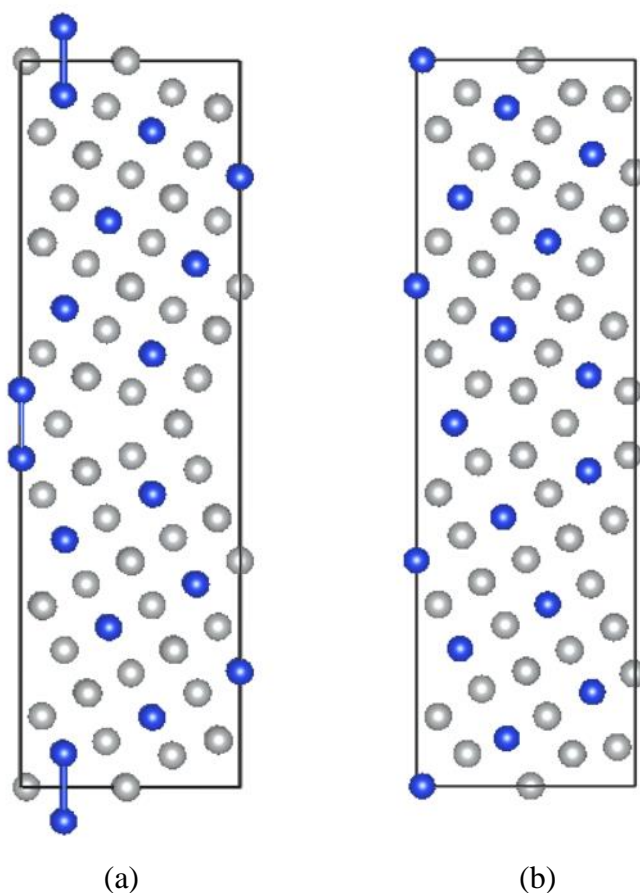
Poté se může začít s vytvořením útvarů mezi atomy, pro které máme určeny křivky COHP, a určením jejich obsahu.

Zároveň se vypočítá pro všechny COHP obsah pod křivkou, který budeme s plochou srovnávat. V této práci byl obsah spočítán numericky pomocí lichoběžníkové metody.

6.5 Zobrazení jedné mřížky na druhou

Po shromáždění dat ze všech dříve popsanych analýz je možné začít porovnávat stabilní a nestabilní mřížku.

V tuto chvíli nastává otázka, jak porovnávat mřížky, které jsou sice podobné (jak je vidět na obr. 8-1), ale neexistuje jednoznačné zobrazení jedné na druhou.



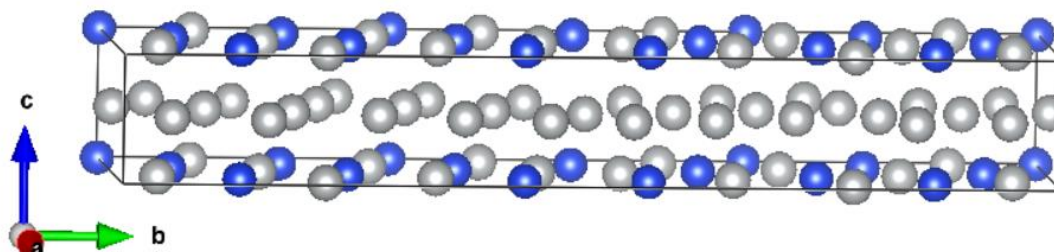
Obrázek 6-27 Porovnání (a) stabilní a (b) nestabilní mřížky [27]

Podobnosti se využije tak, že se najde ke každému atomu stabilní mřížky souřadnicově nejbližší atom mřížky nestabilní. Pro přiřazení to však není dostatečná informace. Může se stát, že jednomu atomu ze stabilní mřížky jsou přiřazeny dva z nestabilní, pro práci je však potřeba zobrazení bijektivní.

Provede se tedy tato operace ještě jednou opačně a k atomům nestabilní mřížky přiřadíme atomy mřížky stabilní. Máme tak dva sady dvojic, ze kterých vybereme jejich

průnik. Tímto způsobem se sice nepodaří přiřadit úplně všechny atomy, ale pro ty důležité tak přiřazení vznikne.

Toto přiřazení však nemusí být finální. Jak je vidět na obr. 8-2, kde je nestabilní mřížka zobrazena z jiného úhlu, skládá se mřížka ze dvou neustále se opakujících vrstev. Může se tedy stát, že přiřazení nebude pro vrstvy se stejnou souřadnicí c (dle obr. 8-2), ale budou posunuté. U jedné mřížky se tedy zkusí vrstvy prohodit a určí se nové přiřazení.



Obrázek 6-28 Pohled na nestabilní mřížku z jiného úhlu [27]

Výrazně nižší průměrná vzdálenost přiřazených atomů vychází v druhém případě s přehozenými vrstvami, což je přiřazení použité i v článku [17].

7 APLIKACE

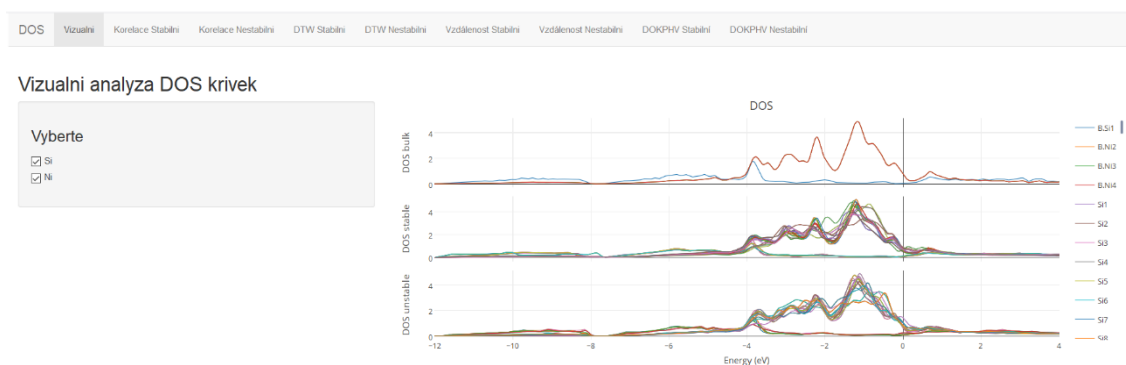
Pro snazší analýzu byly vytvořeny aplikace pomocí knihovny shiny, kde lze jednoduše porovnávat výsledky jednotlivých analýz.

7.1 Křivková analýza DOS

V této aplikaci najdeme křivkové analýzy provedené v kapitole 6.2.

7.1.1 Vizuální kontrola

První část nemá na starosti nic víc, než jednoduše křivky zobrazit a umožnit tak uživateli přibližný náhled na data, se kterými pracuje. (obr. 7-1)



Obrázek 7-1 Aplikace, vizuální kontrola

Tato část obsahuje graf z knihovny plotly, který se skládá ze tří částí – základní mřížka, stabilní mřížka a nestabilní mřížka. Díky interaktivitě, kterou knihovna plotly umožňuje, lze vybírat, které křivky chceme či nechceme zobrazit pomocí klikání v legendě grafu, přiblížit zkoumanou část grafu, stáhnout jako obrázek a mnoho dalšího.

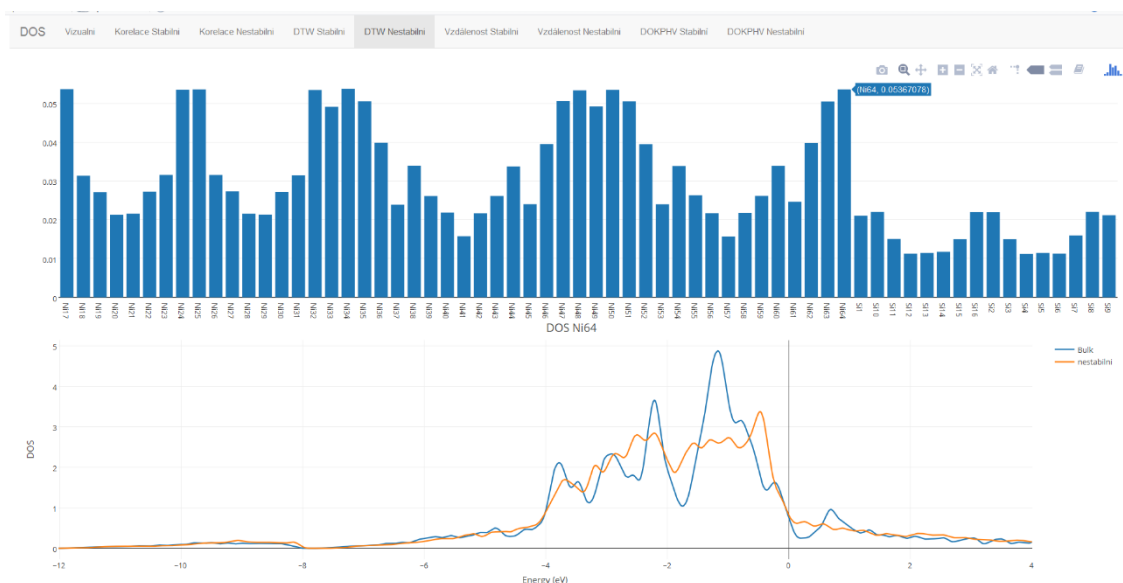
Pro usnadnění výběru je tu možnost z menu vpravo vybrat zobrazení křivek pouze pro Ni nebo Si.

7.1.2 Korelace, vzdálenost, DTW

Části, kde se analyzuje korelace, vzdálenost křivek obyčejná a s DTW, jsou velmi podobné.

Jak je vidět na obr. 7-2, skládá se ze dvou částí. První je sloupcový graf, kde jsou zanesené všechny hodnoty dané analýzy.

V druhém grafu jsou pak zobrazeny křivka pro hodnotu, která byla vybrána uživatelem ve sloupcovém grafu, a odpovídající křivka ze základní mřížky.



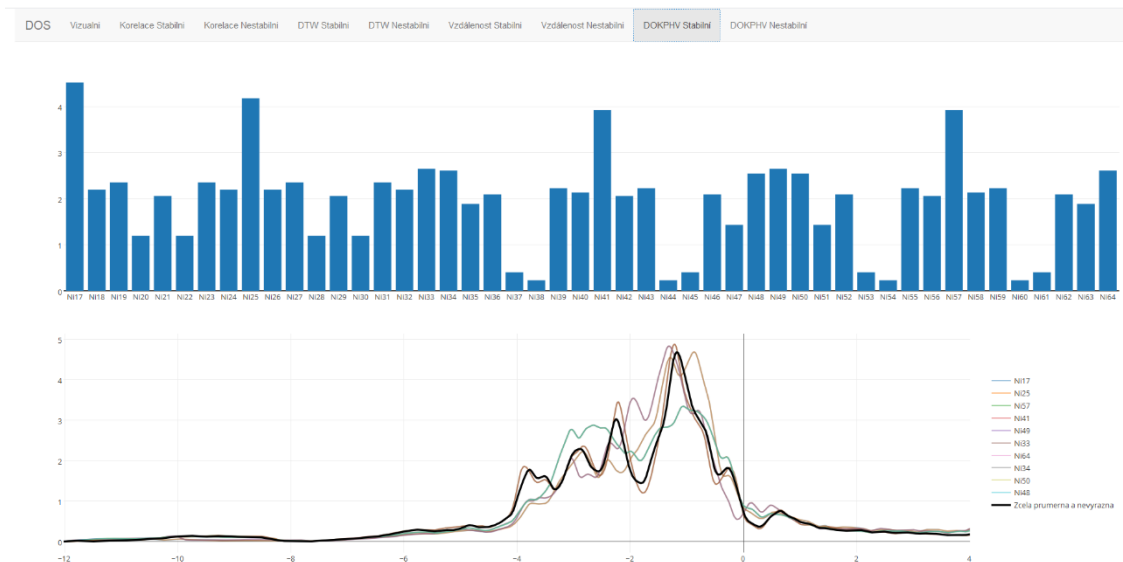
Obrázek 7-2 Aplikace, DTW

7.1.3 Detekce odlehklých křivek pomocí horizontálních vrstevnic

Tato část je jen málo odlišná od těch popsaných v kapitole 7.1.2. Sloupcový graf s vyneseními hodnotami zůstává, změna je však v druhém grafu.

Tentokrát zde nejsou křivky zvolené kliknutím v prvním grafu, ale 10 nejvíce excentrických křivek z daného souboru dat a křivka, který byla vyhodnocena jako průměrná.

Jak tato část vypadá je k nahlédnutí na obr. 7-3.



Obrázek 7-3 Aplikace, DOKPHV

7.2 Křivková analýza COHP

Tato aplikace je velmi podobná té popsané v kapitole 7.1. Nejzásadnější rozdíl je v tom, že zde není záložka Vizuální kontrola (kapitola 7.2.1) z toho jednoduchého důvodu, že křivek je mnohem více, a tak by zobrazení zabralo neúměrné množství času.

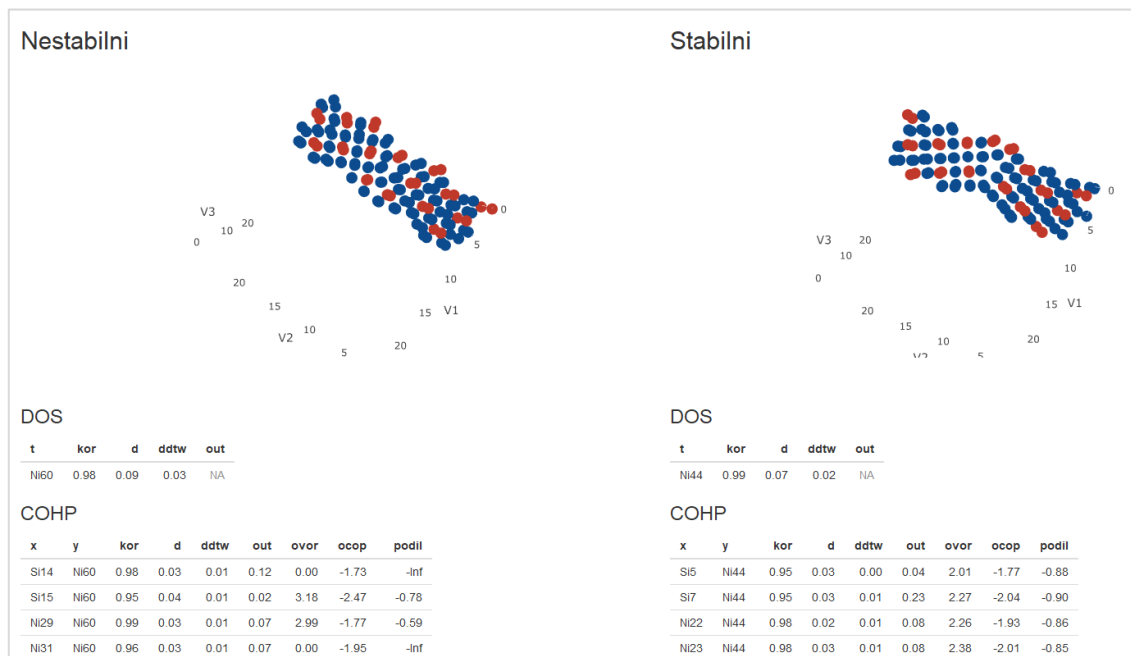
Na obr. 7-4 je vidět, že sloupcový graf obsahuje výrazně více hodnot a stává se mírně nepřehledný.



Obrázek 7-4 Aplikace, COHP

7.3 Zobrazení výsledků analýz

Poslední aplikace slouží k přehlednému zobrazení všech výsledků použitých metod. (obr. 7-5)



Obrázek 7-5 Aplikace, zobrazení výsledků analýz

Aplikaci dominují dva 3D interaktivní grafy z knihovny plotly zobrazující stabilní a nestabilní mřížku. Grafy je možné otáčet, posouvat nebo je přibližovat.

Po kliknutí na libovolný atom se zobrazí veškeré shromážděné informace v tabulkách pod grafy.

8 POROVNÁNÍ MŘÍŽEK

Pro závěrečnou analýzu jsou k dispozici dva druhy informací – ty, které lze přiřadit k jednotlivým atomům, a ty, které lze přiřadit k jejich dvojicím.

Na základě získaných zkušeností bylo rozhodnuto se dále zaměřit především na ty přiřazené dvojicím, tedy informace získané z COHP křivek a Voroného diagramu. Pro připomenutí se jedná o korelaci, minimální vzdálenost, minimální vzdálenost po zarovnání křivek pomocí DTW, excentricitu získanou Detekcí odlehlých hodnot pomocí horizontálních souřadnic, obsah pod křivkou a obsah získaný pomocí Voroného diagramu.

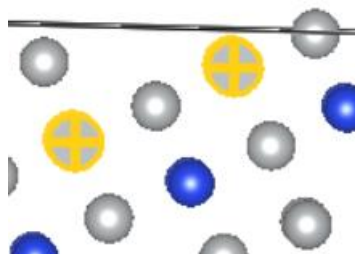
Před začátkem analýzy se musí normalizovat excentricita získaná metodou Detekce odlehlých hodnot pomocí horizontálních souřadnic, protože se jedná o veličinu, která není omezená a v každé analýze může mít jiné hraniční hodnoty. Pro porovnávání bude spíše vypovídající normalizovaná hodnota.

Zavede se ještě jedna hodnota, a to podíl obsahu z Voroného diagramu a obsahu pod křivkou. Dají se u ní předpokládat malé výkyvy, větší odchylka by pak značila nepoměr vazebné síly a vzájemné polohy atomů.

Nyní lze určit rozdíl mezi hodnotami stabilní mřížky a k nim odpovídajícími hodnotami nestabilní mřížky (přiřazení dle kapitoly 6.5) pro všechny metody. Tyto rozdíly jsou vyneseny v grafech na obr. 8-3.

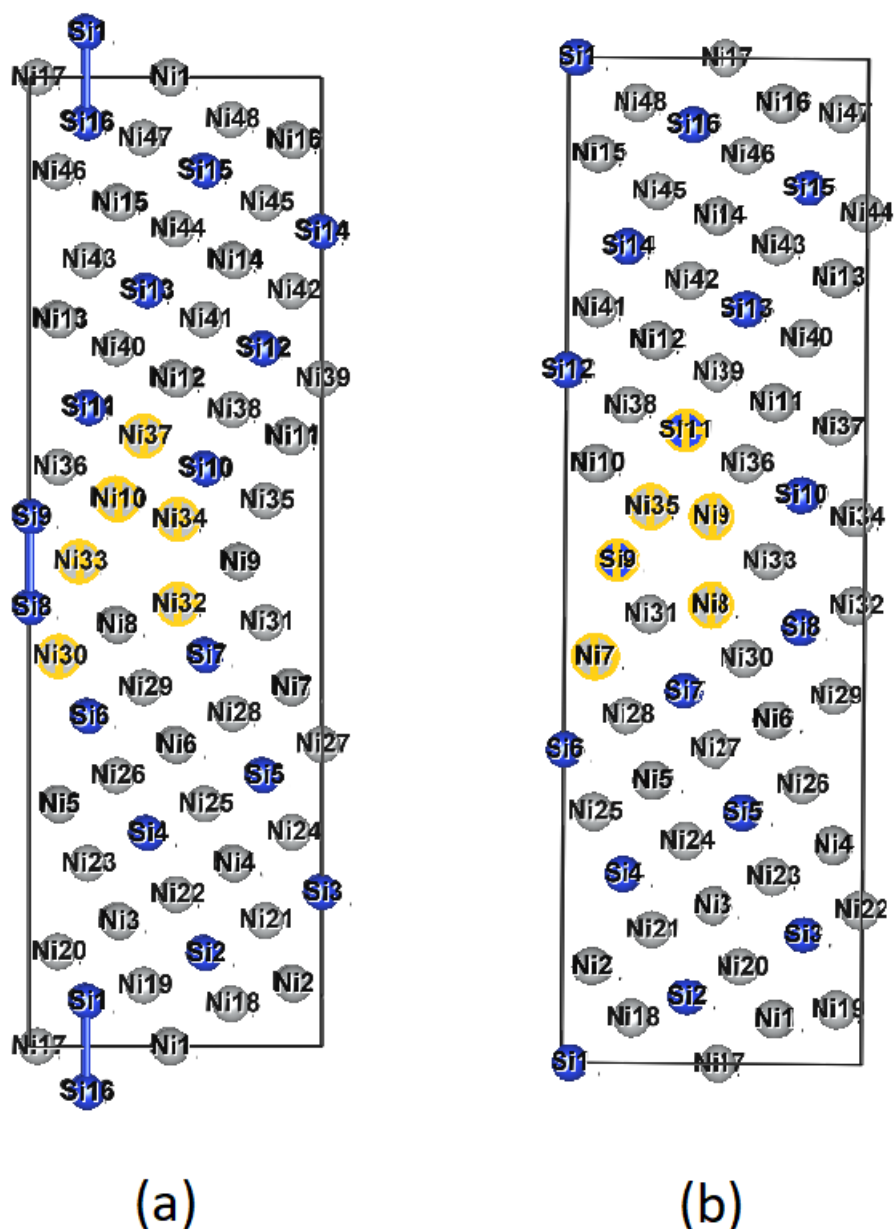
V následujícím textu budou atomy označovány, jak je naznačeno na obr. 8-2. Primárně budou popsány atomy ze stabilní mřížky, v závorce budou uvedeny odpovídající atomy z mřížky nestabilní. Pro usnadnění hledání jsou uvedené atomy označeny na obr. 8-2 žlutě.

Při bližším prozkoumání výsledků rozdílů korelace se ukáže, že největší změny nastávají nejvíce u atomů dál od hranice zrn, které mají vzájemnou polohu, jak je naznačeno na obr. 8-1 (atom, který je na obrázku mezi vyznačenými atomy je v jiné vrstvě).



Obrázek 8-1 příklad atomů s velkou změnou korelace [27]

Asi nejzajímavější rozdíl v minimálních vzdálenostech vznikl mezi atomy Ni32 a Ni34 (Ni8 a Ni9), které se nachází přímo na hranici. Menší, ale přesto zajímavý rozdíl lze zaznamenat i mezi atomy Ni10 a Ni37 (Si11a a Ni35), které taktéž náleží k hranici.



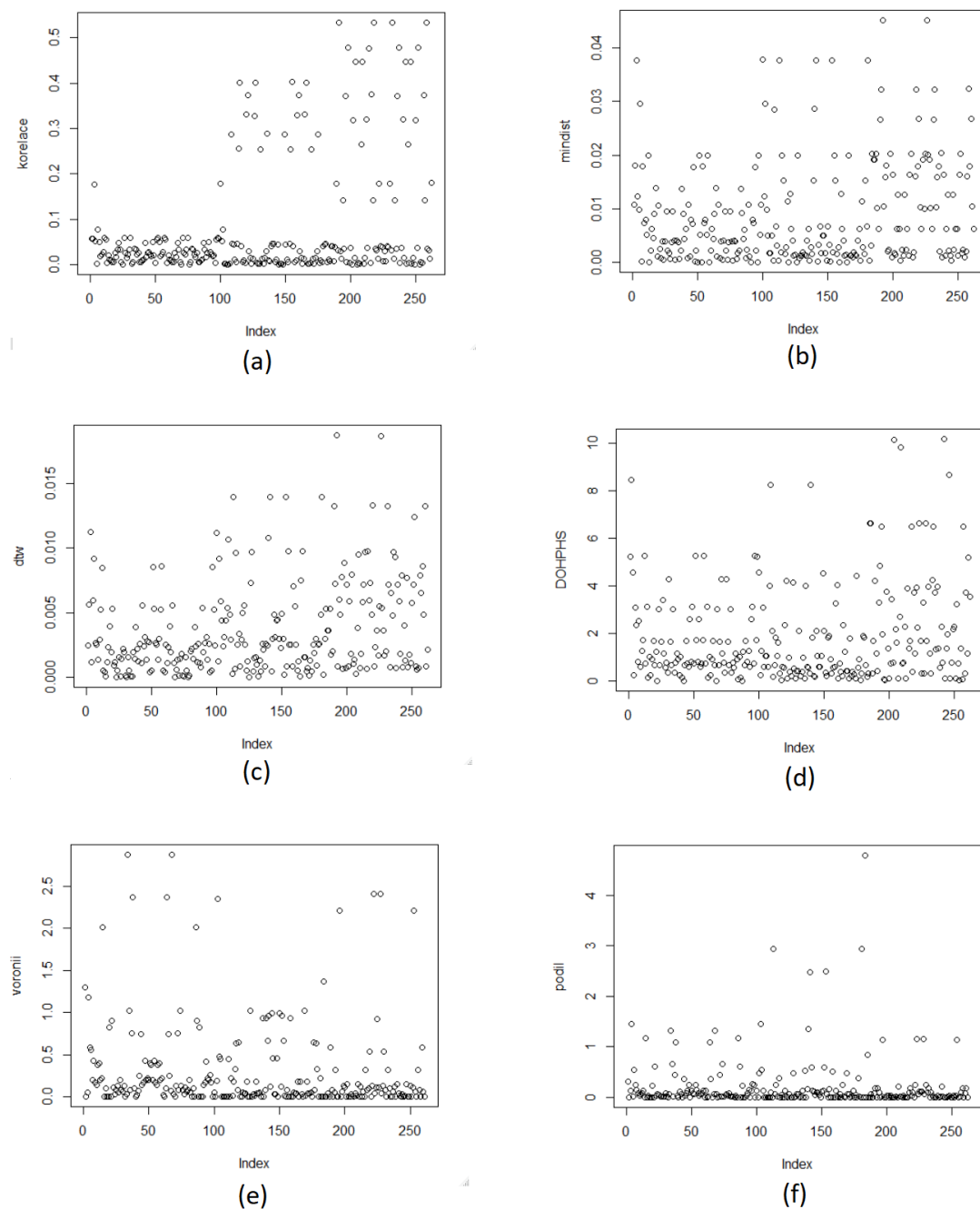
Obrázek 8-2 Očíslování atomů v (a) stabilní mřížce (b) nestabilní mřížce [27]

U metody DTW byly zaznamenány podobné výsledky jako u minimální vzdálenosti. Rozdíly se však zvyrazňují, jak lze vidět na obr. 8-3.

Větších rozdílů v excentricitě si lze povšimnout především u podobného typu dvojic, jako u korelace (obr. 8-1).

Rozdíly v obsahu z Voroného diagramu jsou výrazné například u Ni30 a Ni33 (Si 9 a Ni7) na hranici zrn.

Zajímavý rozdíl pro podíl obsahu z Voroného diagramu a obsahu pod křivkou je vidět například pro atomy Ni30, Ni33 (Si9, Ni7).

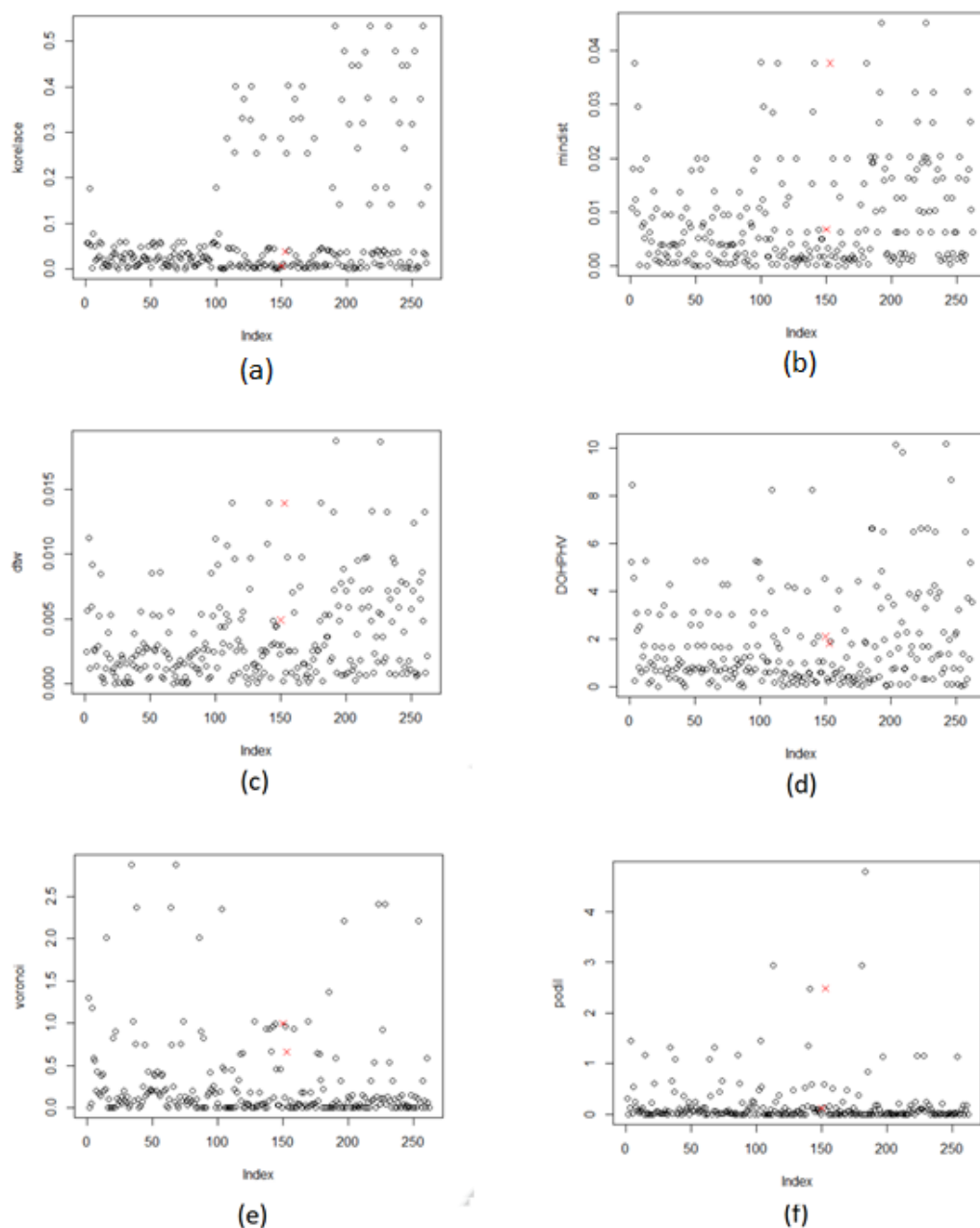


Obrázek 8-3 Porovnání stabilní a nestabilní mřížky, rozdíly hodnot

- (a) Korelace
- (b) Minimální vzdálenost
- (c) Minimální vzdálenost po zarovnání pomocí DTW
- (d) Excentricita získaná pomocí metody DOKPHS
- (e) Obsahy z Voroného diagramu
- (f) Podíl obsahu z Voroného diagramu a obsahu pod křivkou

V článku [17] byl problémový rozdíl identifikován u vazeb Ni10 – Ni37 a Ni10 – Ni33 (Si11 – Ni35 a Si9 – Ni35). První vazbu se podařilo identifikovat pomocí minimální vzdálenosti (s nebo bez DTW), druhou však nikoliv. Na obr. 8-4 jsou označeny červenými křížky hodnoty odpovídající těmto vazbám.

Z grafů je vidět, že všechny metody kromě korelace vyzdvihují alespoň jednu ze dvou problémových vazeb.

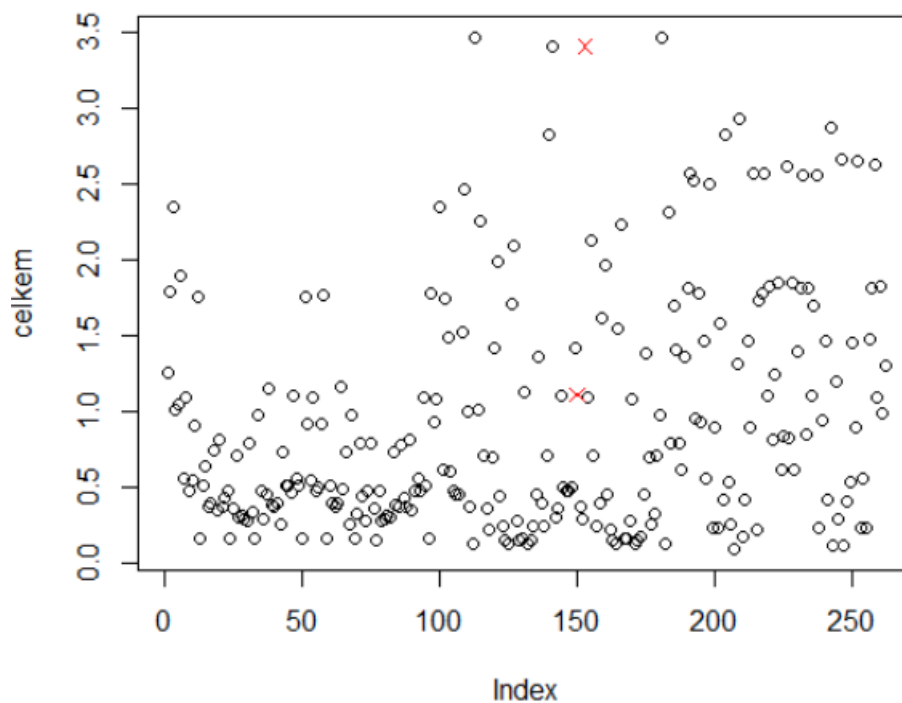


Obrázek 8-4 Vazby pro Ni10 – Ni37 a Ni10 – Ni33

- (a) Korelace
- (b) Minimální vzdálenost
- (c) Minimální vzdálenost po zarovnání pomocí DTW
- (d) Excentricita získaná pomocí metody DOKPHS
- (e) Obsahy z Voroného diagramu
- (f) Podíl obsahu z Voroného diagramu a obsahu pod křivkou

Pro celkové zhodnocení analýzy lze sečíst (zprůměrovat) hodnoty přes všechny metody pro každou vazbu. Hodnoty je však nutné nejdříve normalizovat, protože každá má jiné rozpětí a jejich přínos k celkové hodnotě by pak nebyl rovnoměrný.

Graf na obr. 8-1 znázorňuje součet hodnot. Celková hodnota jednoznačně určuje jednu z problémových vazeb, druhá sice nevykazuje tak výrazný rozdíl, ale stále se vyskytuje v 30 % atomů s nejvyšší hodnotou.



Obrázek 8-1 Celkové skóre analýzy

9 ZÁVĚR

Tato práce byla věnována aplikaci vybraných metod pro rozdílů mezi mřížkou se stabilní hranicí zrna a nestabilní hranicí zrna, u kterých známe křivky DOS a COHP. (více v kapitole 2)

Byly použity metody k porovnávání křivek, konkrétně korelace, minimální vzdálenost, minimální vzdálenost po zarovnání křivek pomocí DTW, excentricita získaná pomocí Detekce odlehlých hodnot pomocí horizontálních souřadnic. (popsány v kapitole 4). Aplikace těchto metod byla poměrně úspěšná, každá byla schopná dobře zaznamenat rozdíl mezi křivkami. (blíže v kapitole 6)

Pro představu, jak vypadá okolí jednotlivých atomů, byl využit Voroného diagram. Z něj byla využita informace o obsahu stěn útvaru obklopující každý atom. Tento obsah lze přiřadit ke dvojici atomů a pak jej porovnávat s obsahem pod křivkou COHP.

Pro snadnější zkoumání dat byly vytvořeny tři aplikace, které jsou popsány v kapitole 7.

Na závěr byla porovnána stabilní a nestabilní mřížka pomocí hodnot získaných výše uvedenými metodami. Rozdíly na hranici zrn byly opravdu nalezeny, ale ukázalo se, že analýzám nelze plně důvěřovat a je dobré je použít spíše jako podporu, díky které lze zjistit, jakým vazbám má cenu se blíže věnovat.

V této závěrečné fázi analýzy se ukázala být nejméně relevantní metoda porovnávání pomocí korelace.

Toto téma má rozhodně potenciál pro další rozvoj. Může se například více využít Voroného diagramu a získat z něj třeba informace o objemu, ploše celého útvaru, nebo jeho topologii. Tyto informace pak lze opět porovnávat a křivkami DOS či COHP.

10 SEZNAM POUŽITÉ LITERATURY

- [1] HAN, Jiawei, Micheline KAMBER a Jian PEI. *Data mining: concepts and techniques*. 3rd ed. Waltham: Morgan Kaufmann, 2012. Morgan Kaufmann series in data management systems. ISBN 978-0-12-381479-1.
- [2] LU, Guang-Hong, Ying ZHANG, Shenghua DENG et al. Origin of intergranular embrittlement of Al alloys induced by Na and Ca segregation: Grain boundary weakening. *Physical Review B*. 2006, **73**(22). DOI: 10.1103/PhysRevB.73.224115. ISSN 1098-0121.
- [3] YAN, Min, M. ŠOB, D. E. LUZZI, V. VITEK, G. J. ACKLAND, M. METHFESSEL a C. O. RODRIGUEZ. Interatomic forces and atomic structure of grain boundaries in copper-bismuth alloys. *Physical Review B*. 1993, **47**(10), 5571-5582. DOI: 10.1103/PhysRevB.47.5571. ISSN 0163-1829.
- [4] BRAITHWAITE, J. Spencer a Peter REZ. Grain boundary impurities in iron. *Acta Materialia*. 2005, **53**(9), 2715-2726. DOI: 10.1016/j.actamat.2005.02.033. ISSN 13596454.
- [5] SUTTON, Adrian P. a R. W. BALLUFFI. *Interfaces in crystalline materials*. New York: Oxford University Press, 1995. ISBN 01-985-1385-2.
- [6] LIU, C.T., E.P. GEORGE a W.C. OLIVER. Grain-boundary fracture and boron effect in Ni₃Si alloys. *Intermetallics*. 1996, **4**(1), 77-83. DOI: 10.1016/0966-9795(95)96901-5. ISSN 09669795.
- [7] VITEK, V. Micromechanisms of intergranular brittle fracture in intermetallic compounds. *Journal de Physique III*. 1991, **1**(6), 1085-1097. DOI: 10.1051/jp3:1991173. ISSN 1155-4320.

- [8] STOLOFF, N.S, C.T LIU a S.C DEEVI. Emerging applications of intermetallics. *Intermetallics*. 2000, **8**(9-11), 1313-1320. DOI: 10.1016/S0966-9795(00)00077-7. ISSN 09669795.
- [9] TAKASUGI, T a O IZUMI. Electronic and structural studies of grain boundary strength and fracture in L12 ordered alloys—I. On binary A3B alloys. *Acta Metallurgica*. 1985, **33**(7), 1247-1258. DOI: 10.1016/0001-6160(85)90236-6. ISSN 00016160.
- [10] TAUB, A.I. a C.L. BRIANT. Composition dependence of ductility in boron-doped, nickel-base L12 alloys. *Acta Metallurgica*. 1987, **35**(7), 1597-1603. DOI: 10.1016/0001-6160(87)90107-6. ISSN 00016160.
- [11] LIU, C.T, C.L WHITE a J.A HORTON. Effect of boron on grain-boundaries in Ni3Al†. *Acta Metallurgica*. 1985, **33**(2), 213-229. DOI: 10.1016/0001-6160(85)90139-7. ISSN 00016160.
- [12] AOKI, Kiyoshi a Osamu IZUMI. Improvement in Room Temperature Ductility of the L1₂ Type Intermetallic Compound Ni₃Al by Boron Addition. *Journal of the Japan Institute of Metals*. 1979, **43**(12), 1190-1196. DOI: 10.2320/jinstmet1952.43.12_1190. ISSN 0021-4876.
- [13] TAKASUGI, T., M. NAGASHIMA a O. IZUMI. Strengthening and ductilization of Ni3Si by the addition of Ti elements. *Acta Metallurgica et Materialia*. 1990, **38**(5), 747-755. DOI: 10.1016/0956-7151(90)90026-D. ISSN 09567151.
- [14] HEATHERLY, L., E.P. GEORGE, C.T. LIU a K.S. KUMAR. An Auger investigation of the grain-boundary chemistry in Ni3(Si,Ti) alloys. *Materials Science and Engineering: A*. 1998, **245**(1), 80-87. DOI: 10.1016/S0921-5093(97)00691-6. ISSN 09215093.

- [15] FRIÁK, Martin, Monika VŠIANSKÁ, David HOLEC a Mojmír ŠOB. Quantum-mechanical study of tensorial elastic and high-temperature thermodynamic properties of grain boundary states in superalloy-phase Ni₃Al. *IOP Conference Series: Materials Science and Engineering*. 2017, **219**. DOI: 10.1088/1757-899X/219/1/012019. ISSN 1757-8981.
- [16] FRIÁK, Martin, Monika VŠIANSKÁ, David HOLEC, Martin ZELENÝ a Mojmír ŠOB. Tensorial elastic properties and stability of interface states associated with $\Sigma 5(210)$ grain boundaries in Ni₃(Al,Si). *Science and Technology of Advanced Materials*. 2017, **18**(1), 273-282. DOI: 10.1080/14686996.2017.1312519. ISSN 1468-6996.
- [17] FRIÁK, Martin, Martin ZELENÝ, Monika VŠIANSKÁ, David HOLEC a Mojmír ŠOB. An Ab Initio Study of Connections between Tensorial Elastic Properties and Chemical Bonds in $\Sigma 5(210)$ Grain Boundaries in Ni₃Si. *Materials*. 2018, **11**(11). DOI: 10.3390/ma11112263. ISSN 1996-1944.
- [18] HAFNER, J. Materials simulations using VASP—a quantum perspective to materials science. *Computer Physics Communications*. 2007, **177**(1-2), 6-13. DOI: 10.1016/j.cpc.2007.02.045. ISSN 00104655.
- [19] DRONSKOWSKI, Richard. *Computational chemistry of solid state materials: a guide for materials scientists, chemists, physicists and others*. Weinheim: Wiley-VCH, 2005. ISBN 35-273-1410-5.
- [20] *The R Project for Statistical Computing* [online]. b.r. [cit. 2019-05-23]. Dostupné z: <https://www.r-project.org/>
- [21] *RStudio* [online]. b.r. [cit. 2019-05-23]. Dostupné z: <https://www.rstudio.com/>
- [22] *Plotly R Open Source Graphing Library* [online]. b.r. [cit. 2019-05-23]. Dostupné z: <https://plot.ly/r/>
- [23] MÜLLER, Meinard. *Information retrieval for music and motion*. New York: Springer, 2007. ISBN 978-3-540-74047-6.
- [24] SMITH, Lindsay I. A tutorial on Principle Components Analysis. *Otago University* [online]. b.r. [cit. 2019-05-23]. Dostupné z: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

- [25] DE MAESSCHALCK, R., D. JOUAN-RIMBAUD a D.L. MASSART. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*. 2000, **50**(1), 1-18. DOI: 10.1016/S0169-7439(99)00047-7. ISSN 01697439.
- [26] AURENHAMMER, Franz. Voronoi diagrams---a survey of a fundamental geometric data structure. *ACM Computing Surveys*. b.r., **23**(3), 345-405. DOI: 10.1145/116873.116880. ISSN 03600300.
- [27] MOMMA, Koichi a Fujio IZUMI. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *Journal of Applied Crystallography*. 2011, **44**(6), 1272-1276. DOI: 10.1107/S0021889811038970. ISSN 0021-8898.

11 SEZNAM PŘÍLOH

1. CD

- Data
 - coh-p-Bulk.csv
 - coh-p-GB-stable.csv
 - coh-p-GB-unstable.csv
 - dos-Bulk.csv
 - dos-GB-stable.csv
 - dos-GB-unstable.csv
- soubory.dat
 - GB-bulk.txt
 - GB-bulk_mat.txt
 - GB-stable.txt
 - GB-stable_mat.txt
 - GB-unstable.txt
 - GB-unstable_mat.txt
- ulozeno
 - vse_dos_nestab.csv
 - vse_dos_stab.csv
 - vse_nestab.csv
 - vse_stab.csv
- app_3D.R
- app_COHP.R
- app_DOS.R
- funkce.R
- krivky_tab.R
- prirazeni.R
- voronoi_aplikovany.R